Dynamic competition and intellectual property rights in a model of product development

Etienne Billette de Villemeur and Richard Ruble and Bruno Versaevel

LEM, Université de Lille, F-59655, France., Emlyon business School, F-69134 (primary affiliation) and CNRS, GATE Lyon Saint-Etienne, Ecully F-69130, France., Emlyon business School, F-69134 (primary affiliation) and CNRS, GATE Lyon Saint-Etienne, Ecully F-69130, France.

September 2017

Dynamic competition and intellectual property rights in a model of product development

Etienne Billette de Villemeur, Richard Ruble and Bruno Versaevel*

August 2, 2018

Abstract

We study innovation timing and socially optimal intellectual property rights (IPRs) when firms facing market uncertainty invest strategically in product development. If demand growth and volatility are high, attrition occurs and IPRs should ensure the cost of imitation attains a lower bound we identify. If demand growth and volatility are low then provided that entry is business-stealing, IPRs should set the cost of imitation high enough to induce preemption, and possibly winner-take-all preemption. Moreover, the welfare achieved with optimal IPRs is then greater with endogenous innovation than if firm roles are predetermined, illustrating the importance of fostering dynamic competition. In extensions we show how firms benefit from open standards, that takeovers have ambiguous welfare effects and that simple licensing schemes are welfare improving.

JEL classification: G31, L13, O33

Keywords: cost of imitation, dynamic competition, patent policy, winner-take-all preemption

*Billette de Villemeur: LEM, Université de Lille, F-59655, France. Ruble and Versaevel: emlyon business School, F-69134 (primary affiliation) and CNRS, GATE Lyon Saint-Etienne, Ecully F-69130, France. Corresponding author: ruble@em-lyon.com. We are grateful to two anonymous referees, Xavier Boutin, Yann Braouezec, Benoît Chevalier-Roignant, David Encaoua, Thilo Meyer-Brandis, Dean Paxson, Régis Renault, Jacco Thijssen, Mihkel Tombak, and Lenos Trigeorgis as well as seminar participants at Paris-Dauphine university, IESEG business school, LMU and TU München, and participants at the 2016 AFSE meeting and Northwestern Annual Conference on Innovation Economics, 2015 International Conference on Real Options and 2014 EARIE meeting for valuable suggestions.

1 Introduction

When developing an invention into a commercial product requires substantial resources, only a few firms may be contenders for positions as first-mover or second entrant in an industry. Along with other key market characteristics like projected growth and volatility, the cost of innovation relative to imitation is then a key determinant of investment strategies and the timing of product introductions. Moreover, there exists substantial variation from one market to another both in dynamic characteristics and in the conditions of imitation, which may themselves hinge on the use that innovating firms make of IPRs. Firms accordingly respond to policy variables, like intellectual property right (IPR) levels, that govern the conditions of innovation and imitation, whose choice by regulators must therefore account for the dynamics of industry investments.

One market in which conditions of imitation and IPRs have played a prominent role is the smartphone market. In the decade following the release of the first iPhone in 2007, a broad interpretation of IPRs allowed Apple to patent numerous aspects of its innovative products and systematically engage in litigation against alleged imitation by its chief rival, Samsung. Many other high-tech firms have pursued similarly aggressive strategies in their own markets (Galasso [15]). In contrast with this, in the growing but still unpredictable market for lithium ion batteries, electric-vehicle technology pioneer Tesla Motors announced in 2014 that it would open its base of several hundreds of approved patents to competitors at no cost, an approach which is not so rare historically in the early stages of new technologies (Bessen [1]). In the pharmaceutical industry, considerable variation can be observed across business segments in both dynamic characteristics and conditions of imitation, which shapes the strategies of the industry's major firms. GlaxoSmithKline for instance announced in 2016 that it would pursue a "graduated approach" to IPRs, under which it would defend its patents in the mature markets of high-income economies but would allow generic manufacturers to produce low-cost copies of its drugs in low-income markets where demand grows faster but can be much more volatile (Phillips [28]). IPRs are known to significantly affect the timing of innovations (drug launches) in this industry, with clear implications for welfare. Several recent empirical investigations have studied innovation timing, such as Wagner and Wakeman [36], who find that the number of patent applications (which can be interpreted as an indirect measure of the cost of "inventing around") for a given drug candidate is positively correlated with the speed of product launch, and Cockburn et al. [7], who find that more extensive patent protection strongly accelerates the introduction of new drugs in a large number of countries. The pharmaceutical sector is an important and well-documented enough industry example that we will continue referring to it throughout the article in a series of detailed footnotes to the main text.

In a model of dynamic competition to develop a new product, we propose to study the effect of the relative costs of innovation and imitation on the investment strategies of firms when their roles as innovator or imitator are endogenous and characterize the regulator's choice of optimal IPR levels. By identifying the role played by the drift and volatility of product market demand on the timing of innovative and imitative investments and hence on economic welfare, we contribute novel insights concerning minimum IPR levels, the necessity of strong IPRs in mature industries, and the role of dynamic competition between firms.

If innovation has positive spillovers for an imitator attrition may arise, and it is all the more likely when demand growth and volatility are high. A socially optimal level of IPRs involves a minimum cost of imitating that we identify. If demand growth and volatility are low, as typically occurs in mature industries, we show that a high level of IPR protection which induces a preemption race constitutes a constrained welfare optimum in a broad range of situations. If demand growth and volatility are sufficiently low, it is even socially desirable to provide innovators with complete protection and have strategic investment take the form of winner-take-all preemption, so that dynamic competition is intense enough that the firms invest at the net present value threshold. The endogeneity of innovation timing plays a key role in establishing these results, and we show that a regulator who did not account for the full range of dynamic competition would run the risk setting too weak a level of IPRs.

To establish these conclusions, we study the exercise of strategic growth options by two initially identical firms pursuing the development of a product for a new market in which they are potential horizontal competitors.[1] Development of innovative and imitative products requires differing stationary levels of irreversible investment which may or may not be related to the exploitation of a single patent, and occurs in a context of market uncertainty as the scale of demand follows a geometric Brownian motion.[2] Both firms independently choose thresholds that determine the timing of their investment in product development, which once performed yields an immediate and perpetual profit flow whose level at any moment depends on the number of active firms. We thus study a real option game,[3] but depart from the standard model by introducing an *ex-post* asymmetry through the differing fixed costs of innovation and imitation that firms face and

[1] Our focus on these industries is therefore complementary to research on cumulative innovation such as Green and Scotchmer [18].

[2] We focus on market uncertainty rather than R&D uncertainty, which has been extensively studied by the patent race literature (see *e.g.* Denicolò [9]).

[3] See Chevalier-Roignant and Trigeorgis [6] for a presentation of these games where firms balance the value of retaining flexibility in the face of uncertainty with the strategic incentive to invest early.

by considering the full range of relative fixed costs. Our model therefore allows just as well for broad IPRs implying a relatively high cost of imitation and preemption between firms as well as for significant spillovers resulting in a comparatively low cost of imitation and attrition between firms.

The imitation cost provides us with a way to parametrize first- and second-mover advantage parsimoniously and to nest within a single framework two important timing games, the war of attrition (Hendricks et al. [20]) and preemption (Fudenberg and Tirole [14]).[4] The timing game we study involves firms choosing investment thresholds, or hurdle rates, that determine stochastic investment times, and has a straightforward normal form. We characterize the unique symmetric equilibrium in investment threshold choices. This analysis provides the foundation for subsequent welfare results which are the main focus of this paper, and we complement it with a more technical discussion of closed-loop strategies in continuous time in the appendix.

After characterizing investment timing and discussing optimal imitation cost levels in a benchmark case with a predetermined sequence of investment decisions instead of dynamic competition (Proposition 1), we derive the symmetric equilibrium timing of innovation with endogenous firm roles for different levels of the imitation cost and identify a critical imitation cost, $\widehat{K}$, which determines whether strategic competition between firms takes the form of attrition or preemption (Proposition 2). For extreme values of the cost of imitation, we find that dynamic competition has the form of a standard timing game. A very low imitation cost induces immediate imitation and leads to a situation of attrition as firms seek to enter second, delaying product introduction.[5] Conversely a very high imitation cost leads to a situation of preemption as firms seek to enter first and enjoy a phase of monopoly profit before imitation occurs. Intermediate values of the imitation cost result in hybrid forms of dynamic competition: a waiting game in which firm investment thresholds are continuously distributed over a disconnected support if the imitation cost is moderately low, and a preemption race in which an attrition phase occurs off the equilibrium path if the imitation cost is moderately high.

Provided that innovation has positive spillovers attrition may occur, and it is more likely if there is a low degree of product market competition or if market growth and volatility are high (Proposition 3). This is because high growth and volatility raise the option value of delaying

[4]The extension of these games to the stochastic case itself presents a number of challenges (Thijssen et al. [34], Steg and Thijssen [32]).

[5]Our model thus encompasses the dynamics described by Scherer (quoted in Fudenberg and Tirole [14]) as "each industry member holding back initiating its R&D effort in the fear that rapid imitation by others will be encouraged, more than wiping out its innovative profits."

investment, eventually compensating for the lost monopoly profit phase if a firm enters second and imitates instead of innovating. A key additional result concerns the optimal balance between first- and second-mover advantage from the standpoint of the industry. Under both attrition and preemption, positional rents are dissipated in the symmetric equilibrium and expected industry value is therefore maximized if the imitation cost attains the critical level $\widehat{K}$ at which there is neither a war of attrition nor a preemption race, so that firms do not compete for positional rents by either unduly waiting or rushing to innovate (Proposition 4).

Because of the tractability of the equilibrium we characterize, we are able to study socially optimal IPR levels if a regulator adjusts the cost incurred by an imitator through either legislative measures or enforcement.[6] With dynamic competition the welfare trade-off associated with raising the imitation cost is more involved than a straight balancing of the incentive to innovate against the deadweight loss of monopoly, as the effect of higher imitation cost on the timing of imitation is ambiguous under attrition. We identify a lower bound on the socially optimal imitation cost (Proposition 5), which must provide sufficient quasi-rents for firms to avoid the Schererian dynamics described in footnote 5 above.

Even if it is generally challenging to draw broad conclusions regarding optimal IPR levels, we are able to show that if the static incentive to imitate is socially excessive, as occurs in a broad range of oligopoly models, an imitation cost that induces preemption is optimal when market growth and volatility are sufficiently low (Proposition 6). The model therefore provides an argument for strong IPRs in such industries based on objective characteristics of market uncertainty. In passing we obtain closed-form expressions for the optimal threshold for innovation and the resulting level of welfare under preemption (Lemma 1), establishing that a limit imitation cost level which results in winner-take-all preemption is socially optimal when there is sufficient discounting. Moreover we provide specific economic circumstances where the optimal imitation cost is consistent either with attrition or preemption, such as a low consumer surplus from innovation or collusion in the

[6]Regulators have a broad array of specific measures at their disposal which affect the cost of imitation. In the pharmaceutical context for example, in the case of biologic-based drugs (*i.e.*, large and complex molecules that are typically produced with recombinant DNA technology), recent US legislation introduced under the Obama Administration guarantees 12 years of data exclusivity. During this period, an imitating firm cannot refer to the originator's clinical data in the regulatory filings for a generic version of an innovative product. In the words of Branstetter, Chatterjee and Higgins [4], (p. 7):

> "With no way to establish bioequivalence, any generic version of a biologic-based drug would have to undergo separate clinical trials to receive FDA approval – a barrier to entry so daunting that no biosimilar has yet been introduced in the U.S. market."

product market (Proposition 7).

These welfare results take on even more relevance when they are compared with the optimal welfare levels derived without dynamic competition, as is often the case in the economic literature on patents. A regulator following this kind of approach could be led to set the level of IPR protection much too low, when in fact competition between firms to innovate plays a vital role and is best incentivized with levels of imitation cost in the preemptive range (Proposition 8).

Finally we discuss several extensions of the model. First, we endogenize the cost of imitation by allowing the innovator to pursue patent protection more aggressively or to make reverse engineering of its product more difficult. A higher baseline cost of imitation reduces the effort exerted by innovators to raise entry barriers, and firms are shown to gain from coordinating *ex-ante* not to introduce subsequent complexity, a policy that may be thought of as an open standard (Proposition 9). We also discuss contracting between innovator and imitator that can take the form either of a takeover or of a license agreement, and show that efficiency always increases in the latter case (Proposition 10).

Our paper is related to early research on innovation incentives and optimal patents, and in particular to Gallini [16] who introduces a cost of imitation that the regulator may use as a policy instrument. We similarly emphasize the role of measures like patent breadth in determining the cost of inventing around an existing innovation, but in contrast with this earlier work we account for the endogenous timing of innovation and thus allow firms to wait before investing rather than assuming that product development occurs when its net present value is positive. Denicolò [9]'s model of optimal IPR protection in a patent race is therefore closer to our work, as it formalizes innovation and imitation as the outcome of a non-cooperative interaction that precedes market competition, though in contrast our model allows for second-mover advantage and attrition, which likely arises in industries with high growth and volatility.

Our work is therefore also related to papers which study the effect of second-mover advantage on investment decisions, often as a result of explicit informational spillovers. Hoppe [23] allows for uncertainty regarding the success of new technology adoption to benefit a rival's innovation decision whereas in Thijssen et al. [33] information regarding the value of a project arrives continuously over time. Femminis and Martini [13] allow for a disclosure lag of random duration before the follower receives the information. In these models, both preemption and attrition can occur depending on the level of spillovers, but the welfare analysis is either based on pure strategy equilibrium or restricted to preemption regimes. We analyze the welfare properties of symmetric mixed strategy equilibrium over a complete range of regimes, providing intuitive analytic results regarding optimal

IPRs that are related to the dynamic properties of demand.

Our welfare results can be related to several papers that compare welfare across two key alternative policy regimes, a strict winner-take-all regime where only the first firm to innovate receives a patent, and a more permissive regime where late investors are allowed to compete with the first before its patent expires. In La Manna, Macleod, and de Meza [25] firms spend a fixed initial amount in R&D that determines a probability of inventing at a future date. Simple cost and demand conditions, such as constant returns to scale and a linear demand, are identified for the permissive regime to be welfare superior. Henry [21] introduces a mechanism whereby a late inventor can share the patent with the innovator within a given time period. When adjusted, together with other policy instruments, this mechanism is socially beneficial under mild conditions, notably with a linear demand and quantity competition. However, in a model where firms incur a flow cost, in Denicolò and Franzoni [11] it is the strict patent regime that is found to be optimal in a broad set of circumstances, in particular when demand is linear, product market competition is weak, and duplication flow costs are large. Our approach is broadly consistent with these contributions, but we characterize an optimal degree of IPRs with the winner-take-all regime occurring as a limit case rather than evaluating a discrete set of regimes. Moreover, the model of investment under market uncertainty allows us to identify determinants of optimal protection that are not considered in this stream of literature related to measurable properties of the dynamics of markets.

Section 2 describes the model. Section 3 studies a benchmark case in which the roles of firms are predetermined. Section 4 characterizes the symmetric equilibrium when firms engage in dynamic competition. Section 5 studies welfare when a regulator determines the cost of imitation. Section 6 discusses two extensions of the model. Section 7 concludes.

2 A model of new product development

This section sets up a model of strategic investment in product development that reflects the characteristic features of innovation and imitation discussed in the introduction. The assumptions concerning the economic environment are presented in Section 2.1 and the continuation payoffs that firms obtain once innovation occurs are presented in Section 2.2.

2.1 Assumptions

Two identical firms engage in dynamic competition to introduce substitutable versions of a novel product in an evolving market. Organizational constraints prevent a firm from developing two variants of the novel product and entry barriers shield both firms from other competitors.

Introduction of the product immediately generates a perpetual baseline profit flow for the firm, whose value is π^M if it operates as a monopoly or π^D if both firms are active, with $0 < 2\pi^D < \pi^M$ which means that imitation reduces industry profits. A firm that introduces its product when no other firm has yet done so is referred to as an *innovator*, and if a firm introduces its product after its rival it is referred to as the *imitator*.

The baseline profit flow is scaled by a measure of market size $Y(t)$, $t \geq 0$. The total flow profits active firms obtain at a given time are either $\pi^M Y(t)$ or $\pi^D Y(t)$. To capture the idea that demand for the new product evolves in a context of market uncertainty, the measure of market size is assumed to follow a geometric Brownian motion $dY(t) = \alpha Y(t)dt + \sigma Y(t)dW(t)$ where $W(t)$ is a standard Wiener process, and α and $\sigma \geq 0$ are the drift and volatility. Both firms have the same discount rate r. In order for the investment problems we study to be economically meaningful we assume that $\alpha < r$.

Product development involves an irreversible investment which encompasses standard setup costs associated with bringing a product to market such as dedicated plant, equipment and marketing expenses as well as the cost of developing the firm's product variant. The fixed costs of innovation and imitation are respectively denoted I and K. While I is assumed to be positive and finite, the extreme cases $K = 0$ and $K = \infty$ are allowed and we are agnostic about the relative magnitudes of I and K.[7] If the second firm can develop an equivalent product completely independently then $K = I$, whereas if there are positive spillovers $K < I$, and with scarce inputs or IPR protection that compels imitators to invent around any intellectual property held by the

[7]Within the pharmaceutical industry for instance, the cost of imitation varies greatly across business segments. The conditions of imitation for drugs are entirely different from those for vaccines. Pharmaceutical firms rely on intellectual property rights in order to increase the costs of imitators for new drugs "which otherwise could be copied more easily than products whose production processes can be kept secret, or for which the time and relative expense needed to copy the invention are much higher" (Scherer and Watal [35], p. 4). If such patent protection is not available, a generic product can be introduced at a much lower fixed cost than incurred by the branded product supplier. However, this ease of imitation is not found in the case of vaccines, which are made from living micro-organisms, and unlike drugs "are not easily reverse-engineered, as the greatest challenges often lie in details of production processes that cannot be inferred from the final product," implying that "there is technically no such thing as a generic vaccine" (Wilson [37], p. 13).

innovator $K > I$ can hold.[8]

We consider an industry that is at an early stage of market development, with $Y(0) \leq (r-\alpha) I/\pi^M$. At such low market size, a monopoly firm would initially prefer to delay investment, and the same holds true for the duopoly firms whose investment decisions we study in the subsequent sections.[9]

Finally, in order to derive our welfare results we suppose that like profits, the baseline consumer surplus flows under monopoly and duopoly, s^M and s^D, with $0 < s^M \leq s^D$, are scaled by the market size parameter $Y(t)$, and that the social discount rate is equal to r. The static welfare gain from imitation, $\left(s^D + 2\pi^D\right) - \left(s^M + \pi^M\right)$, plays a key role in our normative analysis. We assume that this static welfare gain or social incentive for imitation is lower than the private incentive for imitation π^D, implying that $s^D + \pi^D < s^M + \pi^M$. A standard result in industrial economics is that this assumption characterizes a broad range of oligopoly models (Mankiw and Whinston [26]). A set of sufficient conditions on product market competition for it to hold is that entry raise industry output, that it decrease the profit of existing firms and that firms not price below marginal cost.[10]

2.2 Continuation payoffs

Firms obtain continuation payoffs once innovation occurs which depend on their position in the investment sequence. These payoffs are defined for a given value $y = Y(0)$ of the market size process as functions of the market size at which innovation occurs, which is denoted Y and satisfies $Y \geq y$. They thus represent the current values of anticipated rather than instantaneous payoffs,

[8] Imperfect competition in the input market can also lead to asymmetric fixed costs for initially identical firms. Billette de Villemeur et al. [2] show for instance that if the cost of investment is determined endogenously by a monopoly input supplier, the fixed cost is lower for the first firm that invests.

The fixed cost asymmetry we posit can also be contrasted with Pawlina and Kort [27]'s model of dynamic competition in an asymmetric duopoly. Similarly to our equilibrium characterization in Section 4, these authors identify several types of equilibria which are parametrized parsimoniously by the degree of *ex-ante* (rather than *ex-post*) asymmetry.

[9] If this condition is not satisfied, then immediate investment at $Y(0)$ can arise as a solution to the decision problems studied in Sections 3 and 4, and the equilibrium characterizations must be expanded to account for such cases.

[10] The second condition is known as the business-stealing effect, which holds by assumption in our model since $\pi^D < \pi^M$ and characterizes competition between firms producing substitute products in a broad range of industries. In an empirical study of the hypertension drug market for example, Branstetter et al. [5] estimate the impact of imitation (generic entry) on producer profits and on consumer surplus finding only a modest positive effect on social welfare owing to the importance of the business stealing effect (generic sales displace branded sales).

and as in the literature they are denoted F, L and M according to whether a firm invests as a follower, as a leader or if investments are simultaneous. In the terminology of the previous section, the first case corresponds to an imitator whereas the last two correspond to innovator firms, as we rule out imitation with simultaneous investments.

The continuation payoff of a follower is obtained by studying the decision problem of a firm once its rival has innovated. It then holds a growth option on a duopoly market and its optimal policy is to develop the imitative product whenever the market size reaches an exercise threshold we denote by Y^F. Standard arguments (see Appendix A.1) establish that if the current market size is $Y(t) = Y$ the instantaneous value of this option is

$$\begin{aligned} V^D(Y) &= \sup_{\tau \geq t} \mathbb{E}_Y \left[\int_\tau^\infty \pi^D Y(s) e^{-rs} ds - e^{-r\tau} K \right] \\ &= \begin{cases} AY^\beta, & Y < Y^F \\ \frac{\pi^D}{r-\alpha} Y - K, & Y \geq Y^F \end{cases}, \end{aligned} \tag{1}$$

where

$$Y^F := \frac{\beta}{\beta - 1} \frac{r - \alpha}{\pi^D} K$$

is the exercise threshold, β is shorthand for the function of parameters

$$\beta(\alpha, \sigma, r) := \frac{1}{2} - \frac{\alpha}{\sigma^2} + \sqrt{\left(\frac{1}{2} - \frac{\alpha}{\sigma^2} \right)^2 + \frac{2r}{\sigma^2}} \tag{2}$$

and $A := (K/(\beta - 1)) \left[Y^F \right]^{-\beta}$ is a constant. The discounting parameter (2) satisfies $\beta > 1$ and $\partial\beta/\partial\alpha$, $\partial\beta/\partial\sigma$, $-\partial\beta/\partial r < 0$. The expected discounted value at time $t = 0$ of obtaining the duopoly option $V^D(Y)$ at Y when rival innovation occurs is accordingly

$$\begin{aligned} F(Y) &= \mathbb{E}_y \left[\int_{\tau(\max\{Y, Y^F\})}^\infty \pi^D Y(s) e^{-rs} ds - e^{-r\tau\left(\max\left\{Y, Y^F\right\}\right)} K \right] \\ &= \begin{cases} Ay^\beta, & Y < Y^F \\ \left(\frac{\pi^D}{r-\alpha} Y - K \right) \left(\frac{y}{Y} \right)^\beta, & Y \geq Y^F \end{cases} \end{aligned} \tag{3}$$

where in the first line $\tau(Z) := \inf\{t \geq 0 | Y(t) \geq Z\}$ is the stochastic time at which the market size process first hits a threshold Z and Y^F now represents the imitator (or follower) investment threshold.[11]

[11] A standard property of β which is apparent in the continuation payoffs is that the expected discounted value of a monetary unit received at the first hitting time $\tau(Y)$ of threshold $Y \geq y$ is $\mathbb{E}_y e^{-r\tau(Y)} = (y/Y)^\beta$. Note that $\lim_{\sigma \to 0} \beta = r/\alpha$.

Given that the optimal policy of the follower is to invest once Y^F is reached, the leader payoff at time $t = 0$ that a firm receives from innovating first at a threshold Y is

$$\begin{aligned} L(Y) &= \mathbb{E}_y \left[\int_{\tau(Y)}^{\tau\left(\max\left\{Y,Y^F\right\}\right)} \pi^M Y(s) e^{-rs} ds - e^{-r\tau(Y)} I + \int_{\tau(\max\{Y,Y^F\})}^{\infty} \pi^D Y(s) e^{-rs} ds \right] \\ &= \left(\frac{\pi^M}{r-\alpha} Y - I \right) \left(\frac{y}{Y} \right)^\beta - \frac{\pi^M - \pi^D}{r-\alpha} \frac{y^\beta}{\left[\max\left\{Y, Y^F\right\}\right]^{\beta-1}} \end{aligned} \tag{4}$$

(see Appendix A.2). The first summand in (4) corresponds to the discounted value of perpetual monopoly profits from innovation and the second corrects for the anticipated reduction in profit flow stemming from imitation. Moreover, the first summand is a strictly quasiconcave function of Y, which attains its maximum at the standalone monopoly threshold

$$Y^L := \frac{\beta}{\beta-1} \frac{r-\alpha}{\pi^M} I.$$

If both firms invest together finally, at a market size threshold Y, the continuation payoff from simultaneous investments is

$$\begin{aligned} M(Y) &= \mathbb{E}_y \left[\int_{\tau(Y)}^{\infty} \pi^D Y(s) e^{-rs} ds - e^{-r\tau(Y)} I \right] \\ &= \left(\frac{\pi^D}{r-\alpha} Y - I \right) \left(\frac{y}{Y} \right)^\beta . \end{aligned} \tag{5}$$

This payoff is a strictly quasiconcave function of Y which attains its maximum at the threshold

$$Y^M := \frac{\beta}{\beta-1} \frac{r-\alpha}{\pi^D} I,$$

which satisfies $Y^M > Y^L$. The formulation of $M(Y)$ incorporates simultaneous innovation so that the fixed cost is I for both firms, which can be due, say, to a lag for information spillovers to take place ruling out the possibility of any imitation.[12]

3 Predetermined investment sequence

To get a first insight about the dynamics of innovation and imitation in our model, we study investment decisions when firm roles as leader or follower are exogenous. One way this situation

[12] An alternative assumption would be that fixed costs are attributed randomly when investments are simultaneous so the expected fixed cost at the moment of investment is $(I + K)/2$. With this assumption, the first intersection of $L(Y)$ and $M(Y)$ is no longer at Y^F and these two functions no longer overlap if $I \neq K$.

can arise is through certain industrial policies, for example if a former state-run monopolist enjoys priority market access in a deregulated industry. Besides its inherent interest this is a useful case to examine from a theoretical standpoint, first because it leads us to identify a critical imitation cost level, $\widetilde{K}$, which is referred to throughout the rest of our analysis, and second because the equilibrium without dynamic competition that we derive in this section provides a benchmark to assess the effect of dynamic competition further below when we endogenize firm roles.

3.1 Industry equilibrium

Suppose without loss of generality that firm 1 develops its product first whereas firm 2 must wait for firm 1's variant to be introduced before developing its own. Let Y_i, $i \in \{1, 2\}$, denote firm i's investment threshold choice. Firm 2's threshold is therefore constrained to satisfy $Y_2 \geq Y_1$. With these assumptions the equilibrium pattern of investments is found by backward induction as in a standard Stackelberg duopoly model. We therefore first identify firm 2's optimal investment threshold, denoted $Y_2^F(Y_1)$, and then firm 1's optimal forward-looking investment threshold, denoted Y_1^L.

Firm 2 solves the follower problem described at the beginning of Section 2.2. Given the constraint that it faces, its optimal investment threshold is therefore $Y_2^F(Y_1) = \max\left\{Y^F, Y_1\right\}$.

Firm 1 solves the problem $\max_{Y_1 \geq y} L(Y_1)$, where the leader payoff $L(Y_1)$ it obtains is given by (4) which directly incorporates the reaction of firm 2. To solve this problem, observe that the function $L(Y)$ is defined piecewise. Since $L(Y) = M(Y)$ for $Y \geq Y^F$, $L(Y)$ has a local maximum over the range $\left[Y^F, \infty\right)$ at Y^M. If $Y^F > y$, $L(Y)$ is strictly quasiconcave over $\left[y, Y^F\right)$. If $Y^L < Y^F$, which occurs for $K > \left(\pi^D/\pi^M\right) I$, then there is a first local maximum of $L(Y)$ at Y^L. As K increases raising Y^F, the first piece of $L(Y)$ shifts upward continuously (intuitively, a higher threshold Y^F lengthens the leader's profitable monopoly phase) raising the value of the local maximum $L\left(Y^L\right)$, whereas the second subfunction and the value of the other local maximum $M(Y^M)$ are unaffected. For $K = \left(\pi^D/\pi^M\right) I$, $Y^F = Y^L$ so $L(Y^L) = M(Y^L) < M\left(Y^M\right)$, whereas $\lim_{K\to\infty} L(Y^L) = \left(I/\left(\beta-1\right)\right)\left(y/Y^L\right)^\beta > \left(I/\left(\beta-1\right)\right)\left(y/Y^M\right)^\beta = M\left(Y^M\right)$. Therefore, there exists a unique solution in K to the condition $L\left(Y^L\right) = M\left(Y^M\right)$, which we denote by $\widetilde{K}$. Its exact value is

$$\widetilde{K} = \left(\frac{\beta\left(\frac{\pi^M}{\pi^D} - 1\right)}{\left(\frac{\pi^M}{\pi^D}\right)^\beta - 1}\right)^{\frac{1}{\beta-1}} I$$

(see Appendix A.3.1). The critical imitation cost $\widetilde{K}$ allows us to characterize the solution to firm

1's problem: for $K \leq \widetilde{K}$, Y^M is a global maximum of $L(Y)$ whereas for $K \geq \widetilde{K}$, Y^L is a global maximum. For mathematical rigor, if firm 1 is indifferent (for $K = \widetilde{K}$) we assume that it invests at the lower of the two thresholds. Therefore,

$$Y_1^L = \begin{cases} Y^M, & K < \widetilde{K} \\ Y^L, & K \geq \widetilde{K} \end{cases} .$$

With exogenous firm roles generating sequential threshold choices, firms invest at $\left(Y_1^L, Y_2^F\right) = \left(Y^M, Y^M\right)$ (if $K < \widetilde{K}$) or $\left(Y^L, Y^F\right)$ (if $K \geq \widetilde{K}$). In the former case innovation occurs later (since $Y^M > Y^L$) and the outcome involves clustered investments even if firm 2 benefits from a lower fixed cost as an imitator, whereas in the latter case innovation occurs earlier and there is a pattern of diffusion in the industry.

3.2 Efficient imitation cost and IPR levels

We next discuss the normative aspects of investment with exogenous roles from the successive standpoints of the industry and of a regulator, who can in either case set the cost of imitation subject to the constraint that firms thereby determine their investment thresholds in the sequence that is exogenously imposed.

3.2.1 Industry optimum

As seen above, the level of the imitation cost determines the pattern of investments in the industry. This cost in turn depends on technological conditions, but it can also be affected by such measures as *ex-ante* agreements regarding pooling of resources or common standards. It is possible therefore to think of K as a decision variable in certain industries, and to inquire as to what level is optimal from the industry's perspective.

Answering this question amounts to identifying the cost of imitation that maximizes the sum of firm profits $\widetilde{V}_{1+2}(K) := L\left(Y_1^L\right) + F\left(Y_1^L\right)$, that is to solving $\max_{K \in \mathbb{R}_+} \widetilde{V}_{1+2}(K)$. Substituting Y_1^L into the leader and follower continuation payoffs and simplifying, equilibrium industry profit is

$$\widetilde{V}_{1+2}(K) = \begin{cases} \left(\frac{\beta+1}{\beta-1} I - K\right)\left(\frac{y}{Y^M}\right)^\beta, & K < \widetilde{K} \\ \frac{I}{\beta-1}\left(\frac{y}{Y^L}\right)^\beta - \frac{\pi^M - 2\pi^D}{r-\alpha} \frac{y^\beta}{\left[Y^F\right]^{\beta-1}}, & K \geq \widetilde{K} \end{cases} .$$

Industry profit is a decreasing function of K over $\left(0, \widetilde{K}\right)$, discontinuous at $\widetilde{K}$, and increasing over $\left(\widetilde{K}, \infty\right)$ since $\pi^M > 2\pi^D$ and Y^F increases with K. It is straightforward to establish that

$\widetilde{V}_{1+2}(0) < \lim_{K\to\infty} \widetilde{V}_{1+2}(K)$. Therefore the optimal imitation cost from an industry standpoint is $\widetilde{K}^V = \infty$, so that only firm 1 innovates. In such an industry, one would not see development expenses being pooled or common standards being adopted.

3.2.2 Social optimum

The imitation cost can also be affected by policy variables such as the level of IPR protection chosen by regulators. We consider the second-best welfare benchmark here, in which firms are free to select their entry thresholds according to the predetermined investment sequence defined earlier in the section.

The level of social welfare associated with given thresholds Y_1 and Y_2 is

$$\left(\frac{s^M + \pi^M}{r-\alpha} Y_1 - I\right)\left(\frac{y}{Y_1}\right)^\beta + \left(\frac{\left(s^D + 2\pi^D\right) - \left(s^M + \pi^M\right)}{r-\alpha} Y_2 - K\right)\left(\frac{y}{Y_2}\right)^\beta . \tag{6}$$

Substituting for the equilibrium values of Y_1 and Y_2 and simplifying, the welfare second-best with a predetermined investment sequence can be expressed as

$$\widetilde{W}(K) = \begin{cases} \left(\frac{\beta}{\beta-1}\frac{s^D}{\pi^D} I + \frac{\beta+1}{\beta-1} I - K\right)\left(\frac{y}{Y^M}\right)^\beta, & K < \widetilde{K} \\ \left(\beta\frac{s^M}{\pi^M} + 1\right)\frac{I}{\beta-1}\left(\frac{y}{Y^L}\right)^\beta + \left(1 - \beta\frac{\left(s^M+\pi^M\right)-\left(s^D+\pi^D\right)}{\pi^D}\right)\frac{K}{\beta-1}\left(\frac{y}{Y^F}\right)^\beta, & K \geq \widetilde{K} \end{cases} .$$

$\widetilde{W}(K)$ is decreasing for $K < \widetilde{K}$, discontinuous at $\widetilde{K}$, and monotonically decreasing or increasing for $K > \widetilde{K}$ depending on the level of β. For $\beta > \pi^D / \left(\left(s^M + \pi^M\right) - \left(s^D + \pi^D\right)\right)$, $\widetilde{W}(K)$ is increasing in this latter range and it there exists therefore a unique $\widetilde{\beta} > 1$ (see Appendix A.3.2) beyond which the socially optimal imitation cost is $\widetilde{K}^W = \infty$, whereas otherwise $\widetilde{K}^W \in \left\{0, \widetilde{K}\right\}$. Intuitively, large values of β are associated with a significant amount of discounting, in which case the timing of innovation becomes more important for welfare than deadweight loss from lower product market competition, so that it becomes socially advantageous to induce earlier innovation by increasing K to drive Y_1^L to the lower threshold Y^L even if this comes at the expense of monopolization of the product market.

To summarize the results with exogenous firm roles

Proposition 1 *With a predetermined sequence of moves, investments are clustered at Y^M if $K < \widetilde{K}$ and diffused over Y^L and Y^F if $K \geq \widetilde{K}$. The imitation cost $\widetilde{K}^V = \infty$ is efficient for the industry, and socially optimal if there is sufficient discounting (if $\beta \geq \widetilde{\beta}$).*

4 Endogenous innovation and imitation

Suppose the roles of firms as innovator or imitator are not predetermined but instead result from dynamic competition. A non-cooperative timing game therefore determines the sequence of investments. Firms choose innovative investment thresholds that they can update as imitators in a non-strategic continuation phase if rival innovation occurs. Industry dynamics typically consist of a period of inaction before either firm has developed the product over which the strategic interaction plays out, followed possibly by a monopoly phase and a duopoly phase once both firms have developed their own variants of the product. The game is described in Section 4.1, equilibrium in Section 4.2 and the effect of imitation cost on equilibrium thresholds and payoffs is discussed in Section 4.3.

4.1 Firm strategies and payoffs

The strategy of firm i, $i = 1, 2$, consists of a threshold $Y_i \in [y, \infty]$ that triggers its investment when reached for the first time. Strategies are chosen at time zero to determine the stochastic time at which each firm plans to invest, assumed to be a first hitting time, and thus the timing of innovation in the industry. Once innovation occurs any remaining firm revises its investment threshold in the continuation phase which determines when imitation occurs.[13]

To describe the investment game the strategies Y_1 and Y_2 must be mapped into outcomes. For $Y_1 \neq Y_2$ this is straightforward, since one of the firms is the leader and obtains the payoff $L(\min\{Y_1, Y_2\})$ while the other is therefore the follower and obtains the payoff $F(\min\{Y_1, Y_2\})$. For $Y_1 = Y_2$ however, taking the outcome to consist of simultaneous investments with payoffs $M(Y_i)$ for each firm does not correctly represent economic behavior if payoffs satisfy $L(Y_i) \geq F(Y_i) > M(Y_i)$, $i = 1, 2$, *i.e.* if both firms seek to invest whereas it would be optimal for only one to do so. Such cases arise typically in preemption games, and in the discrete time mixed strategy equilibrium that continuous time approximates, investments in fact turn out to be partially coordinated. A standard solution in the literature is to have players use an extension of mixed strategies, known as simple strategies, which appends an intensity function to each player's threshold distributions in order to describe their probabilistic behavior when making simultaneous investment attempts (Fudenberg and Tirole [14], Thijssen et al. [34], see also Appendix B).

[13] This ability of firms to update their thresholds when rival investment occurs is the main difference with Reinganum [29]'s technology adoption game with open-loop strategies in which firms remain committed to their initial threshold choice as outcomes unfold.

For simplicity, we follow an alternative approach in this section which consists in positing a probabilistic tie-breaking rule for such simultaneous investment attempts. This rule is calibrated to yield payoffs that are consistent with the symmetric equilibrium using extended mixed strategies, and notably satisfy the same rent-dissipation property.

If $Y_1 = Y_2 = Y$ we therefore assume that either firm innovates first with probability[14]

$$p(Y) = \begin{cases} \frac{F(Y)-M(Y)}{L(Y)+F(Y)-2M(Y)}, & Y < Y^F \text{ and } L(Y) \geq F(Y) \\ 0, & \text{otherwise} \end{cases} \tag{7}$$

and simultaneous investments accordingly occur with probability $1 - 2p(Y)$.

With these assumptions the payoff of firm i is

$$V(Y_i, Y_{-i}) = \begin{cases} L(Y_i), & Y_i < Y_{-i} \\ p(Y_i) L(Y_i) + p(Y_i) F(Y_i) + (1 - 2p(Y_i)) M(Y_i), & Y_i = Y_{-i} \\ F(Y_{-i}), & Y_i > Y_{-i} \end{cases} \tag{8}$$

and the normal form of the investment game is

$$(\{1,2\}, [y,\infty) \times [y,\infty), (V,V)).$$

A Nash equilibrium of the investment game is a pair of strategies $\left(\widehat{Y}_1, \widehat{Y}_2\right)$ such that $V\left(\widehat{Y}_i, \widehat{Y}_{-i}\right) \geq V\left(Y_i, \widehat{Y}_{-i}\right)$ for all $Y_i \geq y$, $i \in \{1,2\}$. In the next subsection, we describe the unique symmetric Nash equilibrium, which can involve pure or mixed strategies, establishing the formal structure necessary for the subsequent welfare results. Because by construction the payoff function $V(Y_i, Y_{-i})$ encapsulates outcomes of an equilibrium with mixed strategy extensions, the equilibrium obtained in the static investment game of this section is consistent with the continuous time games in closed-loop strategies in the literature. We verify this by studying a dynamic version of the game in the appendix, which does not pose novel difficulties but is more notationally costly.

4.2 Equilibrium

The nature of the timing game that firms play depends on the relative positions of $F(Y)$ and $L(Y)$, which is governed by the level of the imitation cost K.

[14]Observe that in the first line the investment probability solves the condition $pL(Y)+pF(Y)+(1-2p)M(Y) = F(Y)$ so firms are indifferent between the expected payoff from investing at the threshold Y and the follower payoff from postponing investment.

For $K=0$, imitation is free and immediate so $L(Y) < F(Y)$ for all $Y \geq y$. A higher value of K shifts the imitation payoff $F(Y)$ downward and raises the threshold Y^F at which imitation occurs, which lengthens any monopoly phase and raises the payoff $L(Y)$ to innovating first over $\left[y, Y^F\right]$. There exists a unique K that solves $L\left(Y^L\right) = F\left(Y^F\right)$, which we denote $\widehat{K}$. Its exact value is

$$\widehat{K} = \left(\beta \frac{\pi^M}{\pi^D} - (\beta - 1)\right)^{\frac{1}{\beta-1}} \left(\frac{\pi^M}{\pi^D}\right)^{-\frac{\beta}{\beta-1}} I$$

(see Appendix A.4), and $\widehat{K} \in \left(\widetilde{K}, I\right)$. For higher imitation cost levels $K > \widehat{K}$, the set of market size thresholds over which firms strictly prefer innovating first, $\mathcal{P} := \{Y \geq y |\, L(Y) > F(Y)\}$, is non-empty. Intuitively therefore, for lower levels of K (for $K < \widehat{K}$) a firm that imitates always enjoys a second-mover advantage whereas for higher levels of K (for $K > \widehat{K}$) a firm can obtain a first-mover advantage if it innovates first at a threshold $Y_i \in \mathcal{P}$. In the former case, the timing game between firms takes the form of a waiting game, whereas in the latter case it is a preemption game over the set $\mathcal{P}$, which is referred to as the preemption range. If $K > \widehat{K}$, it is useful to define $Y^P := \inf \mathcal{P}$, which is known as the preemption threshold.

Taken together, the imitation cost levels, $\widetilde{K}, \widehat{K}$ and I delimit four continuation payoffs configurations, represented in Figures $1-4$, which reveal the nature of the timing game for different levels of the imitation cost:[15]

$\boldsymbol{K \leq \widetilde{K}}$ (Figure 1). In this case $F(Y) > L(Y)$ for all $Y \geq y$ (since $K < \widehat{K}$), and Y^M is the global maximum of $L(Y)$ (since $K < \widetilde{K}$). Innovating first at any threshold $Y_i < Y^M$ is dominated by innovating at Y^M, and over $\left[Y^M, \infty\right)$ $L(Y)$ $(= M(Y))$ decreases so the investment game constitutes a standard war of attrition.[16]

$\boldsymbol{\widetilde{K} < K < \widehat{K}}$ (Figure 2). In this case $F(Y) > L(Y)$ for all $Y \geq y$ (since $K < \widehat{K}$), but Y^L is the global maximum of $L(Y)$ (since $K > \widetilde{K}$), and $L(Y)$ is not monotone over $\left[Y^L, \infty\right)$ since it has a local maximum at Y^M. Firms therefore engage in a war of attrition which is non-standard. In a symmetric equilibrium innovation thresholds are continuously distributed over a support consisting of the set of market sizes over which $L(Y)$ decreases, $\left[Y^L, \sup\left\{Z < Y^M \middle|\, L(Z) > L(Y^M)\right\}\right] \cup [Y^M, \infty)$ (intuitively it can be useful to think of attrition as occurring over a "decreasing envelope" of $L(Y)$ for $Y \geq Y^L$, given by $\widehat{L}(Y) := \sup\{L(Z) |\, Z \geq Y\}$).

[15] The pivotal cases $K \in \left\{\widetilde{K}, \widehat{K}, I\right\}$ which are not shown are straightforward to obtain from those that are shown by continuity of $F(Y)$ and $L(Y)$ in K.

[16] Figure 1 is drawn assuming $K > (\pi_D/\pi_M)I$ so $Y_F > Y_L$. Otherwise $F(Y)$ is decreasing over (y, ∞) but the key properties described in the text still hold.

$\boldsymbol{\widehat{K} \leq K < I}$ (Figure 3). In this case, the preemption range is non-empty ($K = \widehat{K}$ is a limiting case). Firms therefore race to innovate ahead of one another over any threshold $Y_i \in \mathcal{P}$ and in symmetric equilibrium a single firm innovates at Y^P (or Y^L in the limiting case $K = \widehat{K}$). This preemption race has a non-standard feature, due to the endogenous asymmetry between innovator and imitator fixed costs. Since $\sup \mathcal{P} < Y^F$, off the equilibrium path firms would engage in a war of attrition if the threshold $\sup \mathcal{P}$ is reached and no firm has yet invested.

$\boldsymbol{K \geq I}$ (Figure 4). In this case (which by continuity encompasses a standard preemption game for $K = I$), the innovator payoff $L(Y)$ lies above $F(Y)$ for all $Y \in \left(Y^P, Y^F\right)$, and in symmetric equilibrium a single firm innovates at Y^P.

The investment game has several Nash equilibria involving either pure or mixed strategies. We assume that there are no coordinating mechanisms available so that the firms, being symmetric *ex-ante,* play the same strategies.[17] Moreover, the resulting equilibrium leads to a compelling relationship between industry outcomes and imitation cost. We therefore focus on the unique symmetric equilibrium, consistently with Fudenberg and Tirole [14]'s study of preemption and with the discussion of attrition in Hendricks *et al.* [20].

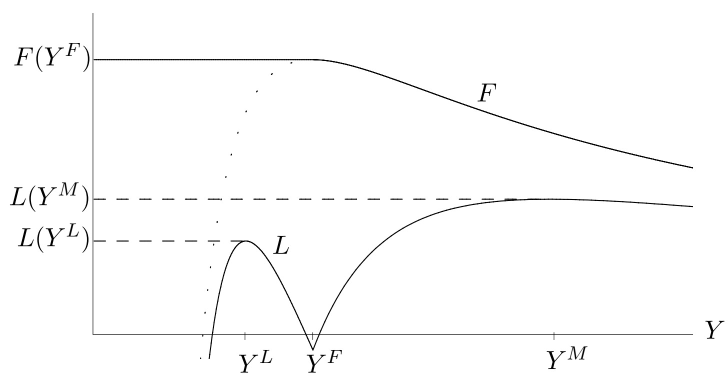

Figure 1 ($K \leq \widetilde{K}$) – Innovation thresholds are distributed over $\left[Y^M, \infty\right)$, and imitation occurs immediately after, for an equilibrium value $\mathbb{E}(V) = L(Y^M)$.

[17]In a similar model of investment with spillovers Hoppe [23] focuses on asymmetric pure strategy equilibria under attrition. Her analysis applies for instance if the firms have multimarket contact that allows investments to be coordinated across markets.

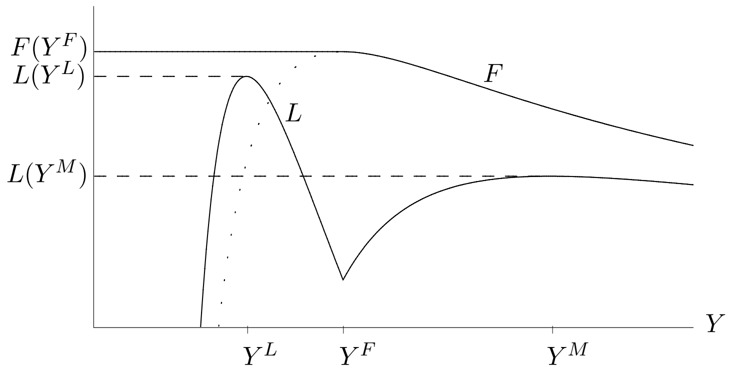

Figure 2 ($\widetilde{K} < K < \widehat{K}$) – Innovation thresholds are distributed over the disconnected support $\left[Y^L, \sup\left\{Z < Y^M \middle| L(Z) > L(Y^M)\right\}\right] \cup \left[Y^M, \infty\right)$, and imitation occurs either at Y^F if the innovation threshold is in $\left[Y^L, \sup\left\{Z < Y^M \middle| L(Z) > L(Y^M)\right\}\right]$, or immediately otherwise, for an equilibrium value $\mathbb{E}(V) = L(Y^L)$.

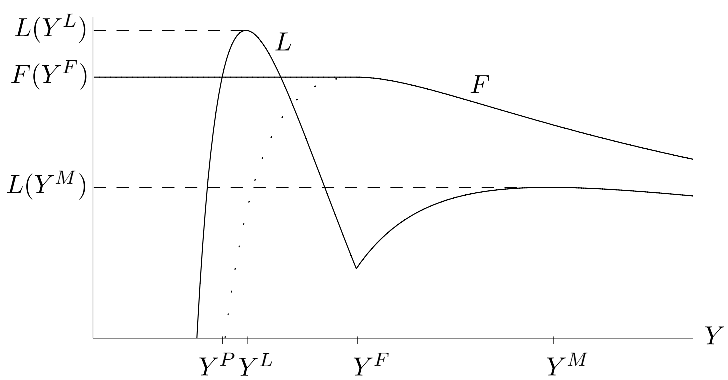

Figure 3 ($\widehat{K} \leq K < I$) – Innovation occurs at Y^P and imitation at Y^F, for an equilibrium value $\mathbb{E}(V) = F(Y^F)$. There is attrition off the equilibrium path if the threshold $\sup \mathcal{P}$ is reached and no firm has yet invested.

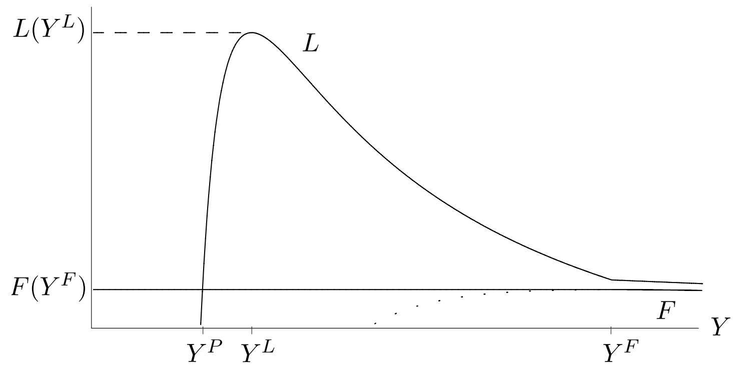

Figure 4 ($K \geq I$) – Innovation occurs at Y^P and imitation at Y^F, for an equilibrium value $\mathbb{E}(V) = F(Y^F)$.

The threshold at which innovation occurs in equilibrium, $\boldsymbol{Y} := \min\left\{\widehat{Y}_1, \widehat{Y}_2\right\}$, is described in the next proposition. This threshold is either a random variable if there is attrition (case (i)), or takes the values Y^P if preemption occurs (case (ii)) or Y^L if neither preemption nor attrition occur (case (iii)) (see Appendix A.4 for proof).

Proposition 2 *In the symmetric equilibrium of the investment game,*

(i) (attrition) if $K < \widehat{K}$ equilibrium is in mixed strategies distributed continuously over $\left[Y^M, \infty\right)$ if $K \leq \widetilde{K}$ and $\left[Y^L, \sup\left\{Z < Y^M \middle| L(Z) > L(Y^M)\right\}\right] \cup \left[Y^M, \infty\right)$ if $\widetilde{K} < K < \widehat{K}$;

(ii) (preemption) if $K > \widehat{K}$ equilibrium is in pure strategies and the innovation threshold is Y^P;

(iii) if $K = \widehat{K}$ equilibrium is in pure strategies and the innovation threshold is Y^L.

Given the equilibrium threshold for innovation, the optimal follower behavior described in Section 2.2 determines the imitation threshold. Letting $\boldsymbol{y}$ denote the realization of $\boldsymbol{Y}$, imitation therefore occurs immediately in case (i) if $\boldsymbol{y} \geq Y^M$, or at Y^F and with a lag in case (i) if $\boldsymbol{y} < Y^M$ or in cases (ii) and (iii).

The pivotal imitation cost level that separates attrition and preemption satisfies $\widehat{K} < I$, so a lower imitation cost is necessary but not sufficient for a second mover advantage to exist and attrition to occur. Too see why, consider an industry in which $K = I$. In such an industry a first-mover that invests optimally earns additional monopoly profits until the market size process hits Y^F. In order for a second-mover advantage to arise and firms to be willing to wait, the cost of imitation must be sufficiently low to compensate the second mover for forgoing this monopoly rent. In practice therefore a lower cost for imitators, which constitutes the most likely situation absent IPRs, does not itself ensure that firms have a second-mover advantage or that they will find it desirable to pursue so-called imitation strategies.

Because the strategic interaction depends on the position of the imitation cost K relative to $\widehat{K}$, the comparative statics of this threshold reveal the effect that the main industry parameters, the intensity of product market competition (π^M/π^D) and the characteristics of the demand process (α and σ), have on the nature of the timing game.

Proposition 3 *The more intense product market competition is and the lower are drift and volatility, the more likely it is that preemption occurs, and conversely for attrition:*

$$\frac{\partial \widehat{K}}{\partial\left(\pi^M/\pi^D\right)} < 0,\ \frac{\partial \widehat{K}}{\partial \alpha}, \frac{\partial \widehat{K}}{\partial \sigma} > 0.$$

To provide intuition for the last inequality, recall that because of market uncertainty there is an option value from waiting. So long as that there is an inherent advantage to imitation ($K < I$), then for large enough levels of drift and volatility such that $K < \widehat{K}$, this option value outweighs any preemptive motive to secure monopoly rents. That is to say, *an attrition regime is more likely in industries with greater trend growth and demand volatility.*[18] This is an important observation in our framework, because it identifies a countervailing force to several mechanisms that are highlighted in the rest of the paper. As the next sections show, firm choices regarding technology or licensing and regulator choices of IPR levels generally make dynamic competition more preemptive. One therefore expects attrition to occur relatively rarely, except in those industries in which market uncertainty is significant.

[18]One example of real-world conditions that might fit such a framework is the following. Pharmaceutical firms face market conditions that impact product introductions and that can vary significantly across geographic areas. In low- and middle-income countries, economic and demographic drivers often imply high demand growth, but political instability can also result in less demand predictability than in high-income economies, and thus discourage the industry from introducing new treatments or preventives. Managers of big pharmaceutical companies are very aware of such market characteristics, and emphasize that although "pharmaceutical markets in key emerging economies, such as China, India, and Brazil, are expanding at rates of more than 12 percent per year (...) uncertain demand, and political and economic instability in some countries have deterred private investors for decades" (Witty [38], pp. 118 and 124).

4.3 Imitation cost and industry outcomes

The equilibrium described in Proposition 2 involves several intuitive relationships between the cost of imitation and the dynamic pattern of investments and firm profitability which we describe here successively.

4.3.1 Thresholds

From the expression of Y^F it directly follows that a higher imitation cost raises the standalone duopoly threshold. However raising Y^F does not by itself imply that imitation is delayed in equilibrium at least in an attrition regime, since innovation and hence imitation occur at a stochastic threshold beyond Y^F with positive probability. The effect of imitation cost on innovation and innovation timing must therefore be studied more carefully.

To begin with, there is an inverse relationship between the cost of imitation and the innovation threshold. Under preemption, it is straightforward to establish that $\partial Y^F/\partial K < 0$. Under attrition, the distribution of innovator entry thresholds is shifted leftward so the innovation threshold $\boldsymbol{Y}$ decreases stochastically.

The relationship between the cost of imitation and the imitation threshold, which is monotone under preemption since $\partial Y^F/\partial K > 0$, is again more involved under attrition. Indeed in the waiting game a higher imitation cost results in a higher imitation threshold if the innovation threshold realization is low so that there is a positive lag between innovation and imitation ($\boldsymbol{y} \leq Y^F$). However if the realization of the innovation threshold is high ($\boldsymbol{y} > Y^F$), imitation occurs right after innovation, so its distribution is accordingly also shifted leftward by an increase in the cost of imitation to the right of Y^F. Therefore a higher imitation cost delays imitation if innovation occurs early, but indirectly hastens imitation if innovation occurs late. A better measure of the speed of imitation is therefore the gap between innovation and imitation thresholds, which under attrition is equal to $\max\left\{Y^F, \boldsymbol{Y}\right\} - \boldsymbol{Y}$ and is nondecreasing in K.

In the model, changes in imitation cost therefore have a monotonic effect on both the innovation threshold and the imitation lag. An increase in imitation cost thus accelerates innovation and delays the arrival of imitation conditional upon innovation having occurred.

4.3.2 Equilibrium payoffs

There also exists a useful equilibrium relationship between imitation cost and industry performance. Observe first that because under attrition and preemption competition between firms to secure second- or first-mover advantage results in the dissipation of positional rents, equilibrium firm values have straightforward expressions. In an attrition regime, the equilibrium firm value is $L\left(Y^L\right) = M\left(Y^M\right)$ over $\left[0, \widetilde{K}\right]$, $L\left(Y^L\right)$ over $\left[\widetilde{K}, \widehat{K}\right]$ and $F\left(Y^F\right)$ over $\left[\widehat{K}, \infty\right)$. Moreover viewed as functions of the cost of imitation, $M\left(Y^M\right)$ is constant, $L\left(Y^L\right)$ is increasing and $F\left(Y^F\right)$ is decreasing over the interiors of the relevant ranges. Industry value is therefore nondecreasing in K up until $\widehat{K}$ and decreasing thereafter, and we can conclude that it is at this level where neither attrition nor preemption occur that industry value is maximized, as firms do not have an incentive to dissipate resources by seeking a positional advantage of either sort.

Proposition 4 *The expected industry value with endogenous firm roles is*

$$\widehat{V}_{1+2}\left(K\right) = \min\left\{F\left(Y^F\right), \max\left\{L\left(Y^L\right), M\left(Y^M\right)\right\}\right\}$$

which is quasiconcave, constant over $\left(0, \widetilde{K}\right)$ *and has a unique maximum at* $\widehat{K}^V = \widehat{K}$.

In economic terms, Proposition 4 establishes that starting from a zero imitation cost, firms benefit *ex-ante* from raising the fixed cost of imitation above $\widetilde{K}$ so as to shield an innovator with positive probability from instantaneous imitation. Moreover, and despite possibly wasteful duplicative fixed costs *ex-post*, raising imitation cost for the second firm even further to $\widehat{K}$ is beneficial for the industry when the endogenous timing of investments is accounted for.

Aside from providing an intuitive characterization of industry value, Proposition 4 is instrumental in the next section in establishing several of the welfare results.

5 Normative analysis

The previous section showed how the nature of the timing game (attrition or preemption) and the dynamics of innovation and imitation are related to the cost of imitation. The social welfare generated by the innovative and imitative products is therefore determined by regulatory choices which affect imitation cost. In this section, we seek to identify socially optimal levels of imitation cost in a second-best framework where regulators set the cost of imitation while firms freely determine the timing of their investments. At first glance this question might seem to just involve a classic trade-off between the reward of innovation and the dynamic deadweight loss from monopoly, since a higher imitation cost *prima facie* raises the optimal threshold for imitation. However as seen in Section 4.3.1, imitation does not necessarily occur at an optimal threshold so an increase in imitation cost has an ambiguous effect on the timing of imitation in an attrition regime, and the optimal imitation cost therefore needs to be studied more carefully.

For our welfare analysis, it is useful to express the social welfare function (6) in terms of producer surplus and the consumer surpluses due to innovation and imitation. Its value in a free entry equilibrium is

$$\begin{aligned}\widehat{W}(K) = 2\min\left\{F\left(Y^F\right), \max\left\{L\left(Y^L\right), M\left(Y^M\right)\right\}\right\} \\ + \frac{s^M y^\beta}{r-\alpha}\mathbb{E}_y \boldsymbol{Y}^{-(\beta-1)} + \frac{\left(s^D - s^M\right) y^\beta}{r-\alpha}\mathbb{E}_y\left[\max\left\{\boldsymbol{Y}, Y^F\right\}\right]^{-(\beta-1)}. \quad (9)\end{aligned}$$

The first summand in (9) is expected industry value, which by Proposition 4 is constant for $K < \widetilde{K}$ and strictly quasiconcave with a maximum at $\widehat{K}$. The second term is the expected consumer surplus due to innovative investment. This term is monotonically increasing in K since a higher imitation cost shifts the distribution of innovation thresholds leftward. The third term is the expected consumer surplus due to the imitator's entry whose relationship with K is ambiguous under attrition, since it is the lag between innovation and imitation and not the timing of imitation itself that increases monotonically with K.

The function $W(K)$ does not have a closed-form expression over its entire range. However its value over $\left[0, \widehat{K}\right]$ has an intuitive bound and there is a semi-closed form over $\left[\widehat{K}, \infty\right)$ whose maximum value we are able to calculate, so several properties of the social optimum can be derived (See appendix for proofs of the following propositions and lemmas).

First, the socially optimal imitation cost has a positive lower bound:

Proposition 5 *If the order of investments is endogenous, the socially optimal cost of imitation satisfies $\widehat{K}^W \geq \widetilde{K}$.*

The intuition for this result is straightforward. The first term (producer surplus) in the welfare function is constant over $\left[0, \widetilde{K}\right]$ (by Proposition 4) and the second term (consumer surplus from innovation) is increasing since a higher imitation cost accelerates innovation, so only the third term (consumer surplus from imitation) requires more careful consideration. However if $K \leq \widetilde{K}$ the innovation threshold is distributed over $\left[Y^M, \infty\right)$, and imitation occurs immediately after innovation. Within this range an increase in imitation cost indirectly accelerates imitation, which unambiguously increases the consumer surplus from imitation.

It is therefore never optimal to set the cost of imitation at zero. Rather, it must be sufficiently high so that an innovator expects a phase of monopoly profits with positive probability provided that he innovates at a low enough threshold. Conversely a firm that "wins" the timing game by being more patient than its rival should accordingly pay a minimum price to develop its imitative product, so that the industry avoids the Schererian dynamics described in the introduction. Even if one adheres to the view that IPRs should be abolished altogether (see Boldrin and Levine [3]), it is nevertheless important to ascertain that the cost of imitation meets such a threshold.[19]

Second, the positive lower bound of the socially optimal imitation cost can be tightened further if an intuitive condition is met and there is sufficient discounting. The next proposition thus provides a rigorous foundation for strong IPRs based on their dynamic characteristics, under the assumption that the static private imitation incentive is socially excessive ($\pi^D > \left(s^D + 2\pi^D\right) - \left(s^M + \pi^M\right)$) which characterizes many standard oligopoly models (see Section 2.1).

Proposition 6 *If the order of investments is endogenous, the socially optimal cost of imitation satisfies* $\widehat{K}^W \geq \widehat{K}$ *provided that the static private entry incentive is socially excessive and* β *is sufficiently large.*

The proposition establishes that the drift and volatility of market size, through their effect on the discounting parameter β, play a key role in identifying which type of dynamic competition is socially optimal. Specifically, it is in those industries for which drift and volatility are not too large that IPR protection should be set sufficiently high for competition between firms to be preemptive, whereas the issue of optimal IPR levels remains an open question if the drift and volatility are significant.[20]

The proposition is established by means of two lemmas.

[19] In practice the lower bound $\widetilde{K}$ can be used to assess initiatives like those in the pharmaceutical industry to reduce the relative cost of imitation and encourage generic competition. In low- and middle-income countries, often characterized by a rapidly expanding and highly uncertain demand which makes attrition more likely (Proposition 3), optimal social welfare in the local market may indeed involve low IPRs and attrition but in all cases requires that a sufficient level of protection be maintained so that there remains a window of market sizes in which an innovator entering sufficiently early is incentivized by a period of monopoly profits.

For example, in order to increase access to antiretroviral drugs to treat HIV infection in the developing world, over the last decades political mobilization has facilitated the production of generic versions of the medicines patented in developed countries (Hoen et al. [22]), which is consistent with our analysis above.

[20] For instance for orphan drugs and rare disease development, the U.S. Food and Drug Administration enacted an enhanced form of IPR protection (Orphan Drug Exclusivity) together with a tax credit that lowers the costs of clinical trials (Grabowski et al. [17]). To the extent that such markets are characterized by low growth and volatility, our analysis offers theoretical support to such regulatory measures.

Lemma 1 *$\widehat{W}(K)$ has a unique maximum over $\left[\widehat{K},\infty\right]$ and there exists $\widehat{\beta} > 1$ such that the socially optimal innovation threshold over this range is*

$$\widehat{Y}_P^W = \begin{cases} \frac{1}{1+\psi} Y^L, & \beta < \widehat{\beta} \\ \frac{r-\alpha}{\pi^M} I, & \beta \geq \widehat{\beta} \end{cases}, \text{ where } \psi := \frac{\frac{s^M}{\pi^M}\left(\frac{\pi^M}{\pi^D} - \frac{\beta-1}{\beta}\right)}{\frac{2}{\beta} + \frac{s^D - s^M}{\pi^D}}.$$

Although we do not have an expression for the socially optimal imitation cost itself, the lemma gives an exact expression for the optimal innovation threshold if firms play a game of preemption (even though preemption thresholds do not themselves have a closed form generally). This optimal innovation threshold lies between the break-even threshold $(r-\alpha) I/\pi^M$ and the standalone monopoly threshold Y^L, and since ψ is increasing in β, when β is large enough the optimum is a corner solution that involves setting an arbitrarily high imitation cost $\widehat{K}^W = \infty$. Imitation then never occurs and instead firms race to enter in winner-take-all preemption, the timing of the monopoly innovation having been driven to the competitive level by the threat of potential entry.

The exact form of the preemption threshold given in Lemma 1 allows the local maximum of $\widehat{W}(K)$ over $\left[\widehat{K},\infty\right]$ to be evaluated. There is no corresponding expression for the maximum value of $\widehat{W}(K)$ over $\left[0,\widehat{K}\right]$, but if the static private entry incentive is socially excessive then the third term in the welfare function has an intuitive bound involving the profits of the imitating firm, and the values of social welfare over the two different ranges can be compared, establishing the proposition.

Lemma 2 *If the static private imitation incentive is socially excessive, then $\widehat{K}^W \geq \widehat{K}$ if*

$$\frac{s^M}{\pi^M} \geq \Omega(\beta), \text{ where } \Omega(\beta) := \frac{3}{2\beta\left(\left(\frac{\beta}{\beta-1}\right)^{\beta-1} - 1\right)}.$$

The function $\Omega(\beta)$ is decreasing in β with $\lim_{\beta\to\infty} \Omega(\beta) = 0$, so Lemma 2 establishes that with sufficient discounting (*i.e.* sufficiently low drift and volatility) the optimal level of social welfare lies in the range over which dynamic competition is preemptive.

The previous two lemmas provide sufficient conditions, both for high IPRs to be socially optimal ($s^M/\pi^M \geq \Omega(\beta)$) and, provided that this is the case, for optimal IPRs to result in winner-take-all preemption ($\beta \geq \widehat{\beta}$).[21] For several standard product market specifications that

[21] It is useful for the examples in the text to note that $\widehat{\beta}$ is the upper root of

$$\beta^2 \frac{s^M}{\pi^M}\left(\frac{\pi^M}{\pi^D} - 1\right) + \beta\left(2\frac{s^M}{\pi^M} - \frac{s^D}{\pi^D}\right) - \frac{s^M}{\pi^M} - 2 = 0$$

satisfy the excess static private entry incentive restriction in the proposition, these conditions are straightforward to verify.

Example 1 *(linear demand)* Suppose that the product market is characterized by a linear inverse demand $P = A - BQ$, $A, B > 0$ and that firms have constant unit variable cost c. Then after normalizing by $(A-c)^2/B$, product market outcomes are $\left(s^M, s^D, \pi^M, \pi^D\right) = (1/8, 2/9, 1/4, 1/9)$. Solving $s^M/\pi^M = 0.5 = \Omega(\beta)$ numerically gives the threshold for $\widehat{K}^W \geq \widehat{K}$ as $\beta \approx 2.5692$. With these values, $\widehat{\beta}$ is the upper root of $5\beta^2 - 8\beta - 20 = 0$ which gives $\widehat{\beta} \approx 2.9541$ as the threshold for winner-take-all preemption.

Example 2 *(isoelastic demand)* Suppose that the product market is characterized by an isoelastic inverse demand $P = AQ^{-1/\varepsilon}$, $A > 0, \varepsilon > 1$, and that firms have constant unit variable cost c. Then product market outcomes satisfy $s^M/\pi^M = \varepsilon/(\varepsilon-1)$, $s^D/\pi^D = 4\varepsilon/(\varepsilon-1)$ and $\pi^M/\pi^D = 4\left(2(\varepsilon-1)/(2\varepsilon-1)\right)^{\varepsilon-1}$. The threshold for the condition $s^M/\pi^M = \varepsilon/(\varepsilon-1) = \Omega(\beta)$ to hold for all ε is $\beta \approx 1.7201$. With these values, $\widehat{\beta}$ is the upper root of

$$\beta^2\left(4\left(\frac{2(\varepsilon-1)}{2\varepsilon-1}\right)^{\varepsilon-1} - 1\right) - 2\beta - \left(3 - \frac{2}{\varepsilon}\right) = 0,$$

which gives

$$\widehat{\beta} = \frac{1 + \sqrt{1 + \left(3 - \frac{2}{\varepsilon}\right)\left(4\left(\frac{2(\varepsilon-1)}{2\varepsilon-1}\right)^{\varepsilon-1} - 1\right)}}{4\left(\frac{2(\varepsilon-1)}{2\varepsilon-1}\right)^{\varepsilon-1} - 1}.$$

This is an increasing function of ε, with $\lim_{\varepsilon\to\infty}\widehat{\beta} = \left(1 + \sqrt{1 + 3\left(4\mathrm{e}^{-1/2} - 1\right)}\right)/\left(4\mathrm{e}^{-1/2} - 1\right) \approx 2.3122$, which therefore gives the threshold for winner-take-all preemption in the social optimum for all ε.

Further welfare results can be obtained in specific cases:

Proposition 7 *If the order of investments is endogenous then*

(i) if the consumer surplus from innovation is sufficiently small ($s^M \approx 0$), then $\widetilde{K} < \widehat{K}^W < \widehat{K}$;

(ii) if there is collusion in the product market ($s^D + 2\pi^D = s^M + \pi^M$), then $\widehat{K}^W \geq \widehat{K}$.

Part (i) complements the previous results of the section by showing that there exist conditions under which values of the imitation cost in the attrition range constitute a social optimum. This is not obvious *a priori* since there is no closed form expression for $\widehat{W}(K)$ over this range, but it

(see Appendix A.7).

is nevertheless possible to show that $\lim_{K \to \widehat{K}^-} \widehat{W}'(K) < 0$. Intuitively, in this case the innovator does not contribute measurably to the consumer surplus and producer surplus is locally insensitive to imitation cost at $\widehat{K}$. Therefore, decreasing imitation cost from $\widehat{K}$ incentivizes imitation and improves welfare. Part (ii) reflects the opposite situation, where imitator entry does not affect consumer surplus because of collusion in the product market. Then only innovation contributes to social welfare, and the role of imitation cost is to incentivize innovation in a preemption regime.

We complete our analysis by comparing welfare levels we identify with endogenous innovation with those obtained in Section 3 where the investment sequence is predetermined. Provided that $s^M/\pi^M \geq \Omega(\beta)$ we can evaluate the innovation threshold and the optimal welfare level (Lemma 1), and we have shown that for sufficiently large β (for $\beta \geq \max\left\{\widetilde{\beta}, \widehat{\beta}\right\}$) a prohibitive imitation cost is optimal both if roles are predetermined and if roles are endogenous, generating either a contestable monopoly or winner-take-all preemption. A regulator may then compare both of these alternatives, and assess the benefits that stem from inducing dynamic competition. A straightforward calculation establishes that dynamic competition raises welfare in this case, so fostering dynamic competition to innovate first is a valuable policy instrument in such industries. This is not always the case though, and in industries in which a lower imitation cost $\widehat{K}^W \in \left(\widetilde{K}, \widehat{K}\right)$ is optimal (see Proposition 7 for an example), dynamic competition lowers welfare because it induces attrition, delaying the onset of consumer surplus flows without raising firm profits.

Proposition 8 *If β is sufficiently large, it is optimal to induce dynamic competition in the form of winner-take-all preemption.*

Observe moreover that the optimal imitation cost levels are sensitive to whether firm roles are predetermined or endogenous, taking values $\widetilde{K}^W \in \left\{0, \widetilde{K}, \infty\right\}$ in the former case while lying over a connected range $\widehat{K}^W \in \left[\widetilde{K}, \infty\right]$ in the latter. This is a particularly relevant consideration because the patent design literature generally takes the value of innovation as given without endogenizing its timing which can result in a starkly different prescription regarding imitation cost,[22] illustrating the importance of accounting for industry dynamics.

[22] Admittedly with either perfect competition or monopoly instead of duopoly, the timing of innovation follows a straightforward investment rule. As Rockett [30] observes, accordingly "most [models] take the identity of the innovator as given." Denicolò [9] is an exception, but his patent race model does not allow for attrition and second-mover advantage.

6 Extensions

In this section, we discuss how some other aspects of innovative activity fit into our model of endogenous innovation and imitation. One of these is the ability of an innovating firm to raise the imitator's entry barrier by making its product more costly to reverse engineer or by strengthening its patentability, as described in the opening examples of the paper. Another is contracting between the innovator and the imitator, which can consist in a takeover of the rival firm or a technology transfer. From a formal standpoint these different extensions add an intermediate decision to the investment game between the innovative and imitative investment decisions. To the extent that they raise the value of innovating, these extensions favor first-mover advantage and the emergence of preemption regimes, with contrasting implications for imitation timing and welfare.

6.1 Endogenous imitation cost

Suppose that the innovating firm may affect the cost of imitation by varying the amount of either technical or legal protection. In case of technical protection, the cost of reverse engineering can be raised by increasing product complexity. For example an innovating firm can render its product more difficult to disassemble, or even add misleading complexity (Samuelson and Scotchmer [31]). In the case of legal protection, wider patents imply higher costs for inventing around to develop a non-infringing imitation, and firms may decide to pursue patent protection more or less aggressively (Encaoua et al. [12]).

Such choices are incorporated into the model by supposing that when it invests at threshold Y_i, an innovating firm chooses how much additional cost, ρ, to incur in order to raise its rival's imitation cost by an amount $f(\rho)$.[23] The imitation cost increase is instantaneous and the function f is taken to be twice differentiable, increasing and concave with $f(0) = 0$ and $\lim_{\rho\to 0} f'(\rho) = \infty$. The fixed costs of the innovator and imitator are accordingly increasing functions $I(\rho)$ and $K(\rho)$.

Proceeding by backward induction, through $K(\rho)$ the cost-raising effort determines the imitator payoff $F(Y;\rho) := F(Y)|_{K=K(\rho)}$ and standalone threshold $Y^F(\rho) := Y^F\big|_{K=K(\rho)}$. At the moment of innovation therefore, an innovator entering at the threshold Y_i therefore faces the decision problem

$$\max_{\rho\in\mathbb{R}_+} L(Y_i;\rho) := \left(\frac{\pi^M}{r-\alpha}Y_i - I(\rho)\right)\left(\frac{y}{Y_i}\right)^{\beta} - \frac{\pi^M - \pi^D}{r-\alpha}\frac{y^{\beta}}{\left[\max\left\{Y_i, Y^F(\rho)\right\}\right]^{\beta-1}}.$$

[23]See Huisman and Kort [24] for a model of preemption with firms similarly competing on both the timing and magnitude of investment.

Let $\rho^*(Y_i)$ denote the solution to this problem for a given innovation threshold Y_i. At an interior solution $Y^F(\rho^*) > Y_i$ and ρ^* satisfies

$$\beta\left(\frac{\pi^M}{\pi^D} - 1\right) f'(\rho^*) = \left(\frac{Y^F(\rho^*)}{Y_i}\right)^{\beta}.$$

A straightforward comparative static argument establishes that the optimal cost-raising effort is increasing in the investment threshold and decreasing in the baseline imitation cost.[24]

To proceed further we focus on the situation where $K \geq \widehat{K}$ so the dynamic competition is naturally preemptive.[25] Allowing the cost of imitation to be endogenous results in a higher leader payoff $L(Y;\rho^*(Y))$ and a lower follower payoff $F(Y;\rho^*(Y))$ than when this cost is exogenous. This makes the investment game even more preemptive. Since equilibrium payoffs are pegged to the follower value under preemption, firms have a lower expected value in equilibrium. To avoid this penalizing outcome firms would prefer to both commit *ex-ante* not to exert any cost-raising effort if they innovate. One way to achieve such a commitment is by agreeing to an open or common technological standard, a measure frequently observed in early stages of a technology's development and which is not desirable for firms if roles are predetermined (by Proposition 1).

Proposition 9 *If the cost of imitation is endogenous and the investment game is naturally preemptive firms benefit from agreeing ex-ante to a common standard.*

6.2 Takeover and licensing

Contracts ranging from acquisitions and pay-for-delay agreements to joint ventures and licensing contracts typically play an important role in innovation decisions. These arrangements have contrasting effects on investment incentives that can be incorporated into our model. Assume that firms can contract once to transfer either productive assets or technology in exchange for a lump sum transfer, φ, from the innovator firm to the imitator, and that the contract is written by the innovator who holds all the bargaining power.

Because imitation decreases industry profit ($2\pi^D < \pi^M$), it is efficient for the industry if an innovator pays its rival not to subsequently enter the market if it can, by taking over its assets or

[24]The latter property is in line with the situation of biopharmaceutical firms (see footnote 7 above) where greater reliance is placed on patenting in the medications segment in which natural entry barriers are low than in the vaccines segment.

[25]This restriction relates specifically to optimal stopping. For high values of the innovation threshold Y_i, corner solutions $\rho^* = 0$ arise that result in a kink of $L(Y;\rho)$. In such cases the threshold strategies firms are assumed to use needn't be optimal investment policies.

engaging in some equivalent measures.[26] Proceeding by backward induction, in the continuation phase that begins when innovation occurs at a threshold Y_i, expected payoff of the potential imitator is $F(Y_i)$. This continuation payoff constitutes a participation constraint in any contract that the innovator offers. The innovating firm can therefore offer a transfer whose value in time $t = 0$ units satisfies $\varphi^*(Y_i) \in [F(Y_i), F(Y_i) + L(Y_i)|_{K=\infty} - L(Y_i)]$ at the moment when it invests. If the innovator has all the bargaining power ($\varphi^*(Y_i) = F(Y_i)$), its payoff if it takes over the second firm has a straightforward expression

$$L^T(Y) := \left(\frac{\pi^M}{r-\alpha}Y - I\right)\left(\frac{y}{Y}\right)^\beta - F(Y).$$

Then, if $\pi^M/\pi^D \geq \beta + 1$ which occurs if either the efficiency effect $(\pi^M - 2\pi^D)$ is sufficiently strong or in industries with sufficiently high demand growth or volatility, attrition does not occur for any level of K. Otherwise, the follower payoff is increasing in its bargaining power, which raises industry profit under preemption but renders preemption less likely.[27]

Whether the possibility of takeovers runs in the interest of the industry or not depends on the cost of imitation. Under preemption expected profits are pegged to the follower value and therefore increase with takeovers if the follower has some bargaining power, whereas if $K < \widehat{K}$, the industry functions naturally under attrition and firm values are pegged to the leader value. Industry profit then increases if takeovers are allowed, so one would expect an active market for acquisitions to develop in such industries, and all the more so if demand growth and volatility are high.

If a takeover is not possible an innovator must contend with follower entry but can recoup revenue from the imitator's investment through a license fee. Suppose that $K = K_0 + K_1$ where K_0 is an incompressible level of imitation cost reflecting such items as distribution and marketing expenses and K_1 denotes the part of the product development cost that can be eliminated through a technology transfer by firm 1. Licensing does not allow the innovator to push back the moment of imitation, but creates an efficiency gain for the industry due to the elimination of a duplicative expenditure. Supposing that the innovator holds all the bargaining power, it sets the maximum

[26] Pay-for-delay agreements (also called "reverse payment settlements") can arise in the pharmaceutical industry, generally in the context of a patent infringement suit brought by a brand-name company against a generic producer that challenges the innovator's IPRs (see Hemphill [19] and Danzon [8]).

[27] As observed by one referee, an interesting case arises if bargaining power is evenly distributed. The innovator's transfer is then $\varphi^*(Y_i) = F(Y_i) + 0.5(L(Y_i)_{K=\infty} - L(Y_i))$, and the condition in K separating attrition from preemption, which has the form $L(Y_i)_{K=\infty} - \varphi^*(Y_i) = \varphi^*(Y_i)$ with takeovers, reduces to $L(Y_i) = 2F(Y_i)$. The corresponding critical threshold $\widehat{K}^T$ with takeovers therefore satisfies $\widehat{K}^T > \widehat{K}$ in this case.

license fee consistent with the participation constraint at the moment of imitation, $-K_1$. Proceeding by backward induction, the expected revenue from licensing then adds a positive term to the leader payoff which becomes

$$L^L(Y) := \left(\frac{\pi^M}{r-\alpha}Y - I\right)\left(\frac{y}{Y}\right)^\beta - \left(\frac{\pi^M - \pi^D}{r-\alpha}\max\left\{Y, Y^F\right\} - K_I\right)\left(\frac{y}{\max\left\{Y, Y^F\right\}}\right)^\beta.$$

As the leader payoff shifts up to the left of Y^F while leaving the follower payoff function unchanged, the investment game is more preemptive with licensing than with takeovers (as in the case of takeovers however, follower bargaining power dampens this effect and eventually reverses this effect). However, whereas the welfare consequences of takeovers are ambiguous (firms weakly benefit and the consumer surplus from innovation increases because product innovation occurs earlier, but the consumer surplus from imitation is eliminated), the welfare consequences of licensing are unambiguously positive. The visible effect of licensing is the reduction in duplicative of R&D efforts as in Gallini [16], but an additional indirect benefit stems from an increase of consumer surplus that results from the acceleration of innovation.

Thus,

Proposition 10 *With contracting between the innovator and the imitator (i) takeovers are the preferred instrument of an innovator and raise industry profit whereas (ii) licensing is Pareto-improving.*

7 Conclusion

We have sought to model the dynamic allocation of resources to innovation and imitation, explicitly incorporating the interrelated investment decisions under uncertainty of imperfectly competitive firms. As compared with the classic literature on innovation and patents, endogenizing the time at which innovation and imitation occur in the presence of market uncertainty allows us to highlight a novel policy channel, in which IPR levels act upon welfare through their effect on dynamic competition.

The main message that emerges from our analysis is a broadly familiar one, insofar as we find that IPRs must be important enough to provide a sufficient incentive for innovation. But by integrating the theory of investment under uncertainty into the analysis of innovation incentives, we are able to pinpoint the role of specific market characteristics which act as key determinants of investment, and thus to provide a grounding for strong IPRs in circumstances that seem likely to be present in mature industries. In such industries we find that the barriers to imitation should be sufficiently high so as to render dynamic competition between firms preemptive, and if discounting is important enough competition should take the form of a winner-take-all contest. Moreover, a regulator who would attempt to accelerate innovation by designating a market leader would lose out on the social benefits of dynamic competition, as the welfare achieved with optimal IPRs is lower with predetermined roles than with endogenous innovation.

In those industries in which growth and volatility are relatively high on the other hand, which are those most typically associated with vibrant innovation, attrition may be effective in ensuring that the additional benefits of imitation resulting from greater product market competition do not arrive excessively late. Even then some degree of IPR protection can be needed if the cost of imitation is extremely low, in order to ensure that a firm that develops an imitative product as the winner of the attrition game nevertheless pays a high enough entry cost so an industry does not become mired in inefficient dynamics.

In practice, antitrust and industrial policy decisions commonly focus on static product market characteristics. The demand characteristics that we have highlighted, demand growth and volatility, play at least as significant a role in determining the investment incentives and product development decisions, and as such should naturally underlie any determination of optimal IPRs. As we have argued throughout in the footnotes, the policy measures taken in at least one emblematic innovation-intensive industry seem to generally echo this analysis.

References

[1] Bessen J (2014) History backs up Tesla's patent sharing, Harvard Business Review Blog, June 13, available at https://hbr.org/2014/06/history-backs-up-teslas-patent-sharing

[2] Billette de Villemeur E, Ruble R, Versaevel B (2014) Investment timing and vertical relationships, International Journal of Industrial Organization 33:110-123.

[3] Boldrin M, Levine D (2013) The case against patents, Journal of Economic Perspectives 27(1):3-22.

[4] Branstetter L, Chatterjee C, Higgins M (2014) Starving (or fattening) the golden goose?: Generic entry and the incentives for early-stage pharmaceutical innovation, NBER Working Paper Series No. 20532:1-43.

[5] Branstetter L, Chatterjee C, Higgins M (2016) Regulation and welfare: evidence from paragraph IV generic entry in the pharmaceutical industry, RAND Journal of Economics 47(4):857-890.

[6] Chevalier-Roignant B, Trigeorgis L (2011) Competitive Strategy: Options and Games, (Cambridge: MIT Press).

[7] Cockburn I, Lanjouw J, Schankerman M (2016) Patents and the global diffusion of new drugs, American Economic Review 106(1):136-164.

[8] Danzon P (2014) Competition and antitrust cases in the pharmaceutical industry, mimeo.

[9] Denicolò V (1996) Patent races and optimal patent breadth and length, Journal of Industrial Economics 44(3):249-265.

[10] Dixit A, Pindyck R (1994) Investment under Uncertainty, (Princeton: Princeton University Press).

[11] Denicolò V, Franzoni L (2010) On the winner-take-all principle in innovation races, Journal of the European Economic Association 8(5):1133-1158.

[12] Encaoua D, Guellec D, Martinez C (2006) Patent systems for encouraging innovation: Lessons from economic analysis, Research Policy 35(9):1423-1440.

[13] Femminis G, Martini G (2011) Irreversible investment and R&D spillovers in a dynamic duopoly, Journal of Economic Dynamics and Control 35(7):1061-1090.

[14] Fudenberg D, Tirole J (1985) Preemption and rent equalization in the adoption of new technology, Review of Economic Studies 52(3):383-401.

[15] Galasso A (2014) Are Apple's patent wars a marketing strategy?, Harvard Business Review Blog, June 12, available at https://hbr.org/2014/06/are-apples-patent-wars-a-marketing-strategy

[16] Gallini N (1992) Patent policy and costly imitation, RAND Journal of Economics 23(1):52-63.

[17] Grabowski HG, DiMasi JA, Long G, (2015) The roles of patents and research and development incentives in biopharmaceutical innovation, Health Affairs 34(2):302-310.

[18] Green J, Scotchmer S (1995) On the division of profit in sequential innovation, RAND Journal of Economics 26(1):20-33.

[19] Hemphill S (2006) Paying for delay: Pharmaceutical settlements as a regulatory design problem, New York University Law Review 81:1553-1623.

[20] Hendricks K, Weiss A, Wilson C (1988) The war of attrition in continuous time with complete information, International Economic Review 29(4):663-680.

[21] Henry E (2010) Runner-up patents: is monopoly inevitable?, Scandinavian Journal of Economics 112(2):417-440.

[22] Hoen E, Berger J, Calmy A, Moon S (2011) Driving a decade of change: HIV/AIDS, patents and access to medicines for all, Journal of the International AIDS Society 14:15-26.

[23] Hoppe H (2000) Second-mover advantages in the strategic adoption of new technology under uncertainty, International Journal of Industrial Organization 18:315-338.

[24] Huisman K, Kort P (2015) Strategic capacity investment under uncertainty, RAND Journal of Economics 46(2):376-408.

[25] La Manna M, MacLeod R, de Meza D (1989) The case for permissive patents, European Economic Review 33(7):1427-43

[26] Mankiw G, Whinston M (1986) Free entry and social inefficiency, RAND Journal of Economics 17(1):48-58.

[27] Pawlina G, Kort P (2016) Real options in an asymmetric duopoly: Who benefits from your competitive disadvantage? Journal of Economics and Management Strategy, Vol 15(1):1-35.

[28] Phillips T (2016) GSK expands 'means-tested' drug licensing for poorest nations, Intellectual Property Magazine, available at https://www.intellectualpropertymagazine.com/patent/gsk-expands-means-tested-drug-licensing-for-poorest-nations-115896.htm

[29] Reinganum J (1981) On the diffusion of a new technology: a game theoretic approach, Review of Economic Studies 48(3):395-405.

[30] Rockett K (2010) Property rights and invention in Bronwyn H and Rosenberg N (eds) Handbook of the Economics of Innovation (Amsterdam: North-Holland).

[31] Samuelson P, Scotchmer S (2002) The law and economics of reverse engineering, Yale Law Journal 111:1575-1663.

[32] Steg J-H, Thijssen J (2015) Quick or persistent? Strategic investment demanding versatility, Bielefeld: Center for Mathematical Economics working paper 541.

[33] Thijssen J, Huisman K, Kort P (2006) The effects of information on strategic investment and welfare, Economic Theory 28(2):399-424.

[34] Thijssen J, Huisman K, Kort P (2012) Symmetric equilibrium strategies in game theoretic real option models, Journal of Mathematical Economics 48(4):219-225.

[35] Scherer F, Watal J (2002) Post-TRIPS options for access to patented medicines in developing countries, Journal of International Economic Law 5(4):913-939

[36] Wagner S, Wakeman S (2016) What do patent-based measures tell us about product commercialization? Evidence from the pharmaceutical industry, Research Policy 45(5):1091-1102.

[37] Wilson P (2010) Giving developing countries the best shot: An overview of vaccine access and R&D, Oxfam International, 28 pages.

[38] Witty A (2010) New strategies for innovation in global health: A pharmaceutical industry perspective, Health Affairs 30(1):118-126.

A Appendix

A.1 Duopoly option value $V^D(Y)$

Once the rival firm has invested at the market size threshold Y any remaining firm holds a standard growth option (Dixit and Pindyck [10]) whose value is obtained by solving the optimal stopping problem

$$V^D\left(Y\right) = \sup_{\tau \geq t} \mathbb{E}_Y \left(\int_{\tau}^{\infty} \pi^D Y(s) e^{-rs} ds - K e^{-r\tau} \right).$$

From the Hamilton-Jacobi-Bellman equation, the value function $V^D(Y)$ satisfies

$$rV^D\left(Y\right)dt = \mathbb{E}_Y dV^D\left(Y\right)$$

and expanding the right-hand side using Itô's lemma yields the ordinary differential equation that $V^D(Y)$ solves in the continuation region,

$$rV^D(Y) = \alpha Y \left[V^D(Y)\right]' + \frac{1}{2}\sigma^2 Y^2 \left[V^D(Y)\right]'',$$

along with the boundary and smooth pasting conditions

$$V^D(0) = 0$$
$$V^D\left(Y^F\right) = \frac{\pi^D}{r-\alpha}Y^F - K$$
$$V^D\left(Y^F\right) = \frac{\pi^D}{r-\alpha}.$$

The function $A_1 Y^{\beta_1} + A_2 Y^{\beta_2}$ is a candidate solution. The associated fundamental quadratic is $0.5\sigma^2\beta(\beta-1) + \beta\alpha - r = 0$ which has two roots of which only

$$\beta = \frac{1}{2} - \frac{\alpha}{\sigma^2} + \sqrt{\left(\frac{1}{2} - \frac{\alpha}{\sigma^2}\right)^2 + \frac{2r}{\sigma^2}}$$

is positive. Setting $A_2 = 0$ to satisfy the first boundary condition, it follows from the other conditions that

$$Y^F = \frac{\beta}{\beta-1}\frac{r-\alpha}{\pi^D}K$$

and that A_1 is $\left(\pi^D/(\beta(r-\alpha))\right)\left[Y^F\right]^{-(\beta-1)}$ which yields the expression for A in the text.

A.2 Derivation of $L(Y)$ and $M(Y)$

Whereas the expression for the follower payoff (3) is simply the discounted value of $V^D(Y)$, several steps are necessary to obtain (4) from the definition of $L(Y)$:

$$\begin{aligned}
&\mathbb{E}_y\left[\int_{\tau(Y)}^{\tau\left(\max\left\{Y,Y^F\right\}\right)} \pi^M Y(s)e^{-rs}ds - e^{-r\tau(Y)}I + \int_{\tau(\max\{Y,Y^F\})}^{\infty} \pi^D Y(s)e^{-rs}ds\right] \\
&= \mathbb{E}_y\left[\int_{\tau(Y)}^{\infty} \pi^M Y(s)e^{-rs}ds - e^{-r\tau(Y)}I - \int_{\tau(\max\{Y,Y^F\})}^{\infty} \left(\pi^M - \pi^D\right) Y(s)e^{-rs}ds\right] \\
&= \left(\frac{\pi^M}{r-\alpha}Y - I\right)\left(\frac{y}{Y}\right)^\beta - \frac{\pi^M - \pi^D}{r-\alpha}\frac{y^\beta}{\left[\max\left\{Y,Y^F\right\}\right]^{\beta-1}}.
\end{aligned}$$

Finally the derivation of the simultaneous payoff (5) is similar to (4), noting that the flow profit upon investment is π^D instead of π^M and that there is no follow-on entry so the second integral dropped altogether.

A.3 Predetermined investment sequence

A.3.1 Critical imitation cost level $\widetilde{K}$

As stated in the text, $\widetilde{K}$ is defined as the unique solution in K to $L\left(Y^L\right) = M\left(Y^M\right)$. Since $\widetilde{K} \in \left(\left(\pi^D/\pi^M\right) I, \infty\right)$ necessarily and using the definitions (4) and (5), the condition can be expressed as

$$\left(\frac{\pi^M}{r-\alpha}Y^L - I\right)\left(\frac{y}{Y^L}\right)^\beta - \frac{\pi^M - \pi^D}{r-\alpha}Y^F\left(\frac{y}{Y^F}\right)^\beta = \left(\frac{\pi^D}{r-\alpha}Y^M - I\right)\left(\frac{y}{Y^M}\right)^\beta.$$

Substituting for Y^L, Y^F and Y^M and dividing both sides by $\left(I/\left(\beta-1\right)\right)\left(y\left(\beta-1\right)/\beta\left(r-\alpha\right)\right)^\beta$ yields

$$\left(\frac{\pi^M}{I}\right)^\beta - \beta\left(\frac{\pi^M}{\pi^D} - 1\right)\left(\frac{\widetilde{K}}{I}\right)\left(\frac{\pi^D}{\widetilde{K}}\right)^\beta = \left(\frac{\pi^D}{I}\right)^\beta$$

which after rearrangement gives the expression in the text.

A.3.2 Social welfare

Suppose that $\beta > \pi^D/\left(\left(s^M + \pi^M\right) - \left(s^D + \pi^D\right)\right)$ so $\widetilde{W}(K)$ is decreasing over $\left(0, \widetilde{K}\right)$ and increasing over $\left(\widetilde{K}, \infty\right)$. Then to determine the socially optimal imitation cost it is sufficient to compare $\widetilde{W}(0)$ and $\lim_{K\to\infty}\widetilde{W}(K)$. Evaluating, the condition $\widetilde{W}(0) < \lim_{K\to\infty}\widetilde{W}(K)$ is

$$\left(\beta\frac{s^D}{\pi^D} + \beta + 1\right)\frac{I}{\beta-1}\left(\frac{y}{Y^M}\right)^\beta < \left(\beta\frac{s^M}{\pi^M} + 1\right)\frac{I}{\beta-1}\left(\frac{y}{Y^L}\right)^\beta.$$

Substituting for Y^L and Y^M, and dividing both sides by $\left(I/\left(\beta-1\right)\right)\left(y\left(\beta-1\right)\pi^D/\beta\left(r-\alpha\right)\right)^\beta$ yields an equivalent condition after rearrangement,

$$f_{\text{A.3}}\left(\beta\right) := \left(\frac{\pi^M}{\pi^D}\right)^\beta\left(\beta\frac{s^M}{\pi^M} + 1\right) - \left(\beta\frac{s^D}{\pi^D} + \beta + 1\right) > 0.$$

The function on the left-hand side satisfies $f_{\text{A.3}}\left(1\right) = \left(\left(s^M + \pi^M\right) - \left(s^D + 2\pi^D\right)\right)/\pi^D < 0$ and $\lim_{\beta\to\infty} f_{\text{A.3}}\left(\beta\right) = \infty$. Moreover it is strictly convex since

$$f''_{\text{A.3}}\left(\beta\right) = \left(\frac{\pi^M}{\pi^D}\right)^\beta\left(2\frac{s^M}{\pi^M}\ln\frac{\pi^M}{\pi^D} + \left(\beta\frac{s^M}{\pi^M} + 1\right)\left(\ln\frac{\pi^M}{\pi^D}\right)^2\right) > 0.$$

Therefore there exists a unique $\beta_{\text{A.3}} > 1$ such that $f_{\text{A.3}}\left(\beta\right) > 0$ if and only if $\beta > \beta_{\text{A.3}}$. Letting $\widetilde{\beta} := \max\left\{\beta_{\text{A.3}}, \pi^D/\left(\left(s^M + \pi^M\right) - \left(s^D + \pi^D\right)\right)\right\}$, $\widetilde{K}^W = 0$ (if $\beta \le \widetilde{\beta}$) or ∞ (if $\beta \ge \widetilde{\beta}$).

A.4 Proposition 2 (endogenous investment symmetric equilibrium)

A.4.1 Critical imitation cost level $\widehat{K}$

As stated in the text, $\widehat{K}$ is defined as the unique solution in K to $L\left(Y^L\right) = F\left(Y^F\right)$. It is straightforward to check that $\widehat{K} > \widetilde{K}$, since if $K \leq \widetilde{K}$, $L\left(Y^L\right) < L\left(Y^M\right) = M\left(Y^M\right) < F\left(Y^M\right) < F\left(Y^F\right)$, where the last two inequalities respectively hold because $\widetilde{K} < I$ and because Y^F maximizes Y^F. Using the definitions (3) and (4), the condition can be expressed as

$$\left(\frac{\pi^M}{r-\alpha}Y^L - I\right)\left(\frac{y}{Y^L}\right)^\beta - \frac{\pi^M - \pi^D}{r-\alpha}Y^F\left(\frac{y}{Y^F}\right)^\beta = \left(\frac{\pi^D}{r-\alpha}Y^F - \widehat{K}\right)\left(\frac{y}{Y^F}\right)^\beta.$$

Substituting for Y^L and Y^F, and dividing both sides by $(1/(\beta-1))\left(y\left(\beta-1\right)/\beta\left(r-\alpha\right)\right)^\beta$ yields

$$I\left(\frac{\pi^M}{I}\right)^\beta - \beta\left(\frac{\pi^M}{\pi^D} - 1\right)\widehat{K}\left(\frac{\pi^D}{\widehat{K}}\right)^\beta = \widehat{K}\left(\frac{\pi^D}{\widehat{K}}\right)^\beta$$

which after rearrangement gives the expression in the text.

A.4.2 Part (i)

There are two subcases for this part, $K \leq \widetilde{K}$ and $\widetilde{K} < K < \widehat{K}$ (see Figures 1 and 2).

$K \leq \widetilde{K}$ subcase This is a standard war of attrition (in thresholds) over $\left(Y^M, \infty\right)$. Therefore provided that firms do not move with positive probability in $\left[y, Y^M\right]$ the equilibrium distribution follows from the argument in Hendricks et al. [20]. In a nondegenerate (and symmetric) mixed strategy equilibrium, firms randomize investment triggers continuously over $\left[Y^M, \infty\right)$. To derive the equilibrium distribution assume that firm $j \neq i$ randomizes her investment trigger at $t = 0$ according to a cumulative distribution function G. Firm i's expected payoff from investing at a threshold $Y_i \geq Y^M$ is

$$\int_{Y^M}^{\infty} V(Y_i, s)g(s)ds = \int_{Y^M}^{Y_i} F(s)g(s)ds + \left(1 - G(Y_i)\right)M(Y_i).$$

Firm i will randomize if $\partial\left(\int_{Y^M}^{\infty} V(Y_i, s)g(s)ds\right)/\partial Y_i = 0$ over the support, that is if G satisfies $\left[F\left(Y\right) - M(Y)\right]g\left(Y\right) = -M'(Y)\left[1 - G\left(Y\right)\right]$ for all $Y \in \left(Y^M, \infty\right)$. As the same condition holds for firm $-i$, the equilibrium distribution for each firm is

$$G_a\left(Y\right) = 1 - \exp\int_{Y^M}^{Y}\frac{M'(s)}{F\left(s\right) - M(s)}ds$$

and substituting for F and M gives

$$G_a(Y) = \begin{cases} 0, & Y < Y^M \\ 1 - \left(\frac{Y}{Y^M}\right)^{\beta\frac{I}{I-K}} e^{-\beta\frac{I}{I-K}\left(\frac{Y}{Y^M}-1\right)}, & Y \geq Y^M. \end{cases}$$

Then, the distribution of the equilibrium innovation threshold for the industry $\boldsymbol{Y}$ is just that of the minimum of the firm thresholds, $1 - (1 - G_a(\boldsymbol{y}))^2$.

It is claimed in the text (Section 4.3.1) that an increase in imitation cost accelerates innovation. As for $Y > Y^M$,

$$\frac{\partial G_a}{\partial K} = \left(\frac{Y}{Y^M}\right)^{\beta\frac{I}{I-K}} e^{-\beta\frac{I}{I-K}\left(\frac{Y}{Y^M}-1\right)} \beta\frac{I}{(I-K)^2}\left(\frac{Y}{Y^M} - 1 - \ln\frac{Y}{Y^M}\right) > 0$$

(the last bracketed term is positive by the logarithm inequality), the distribution of each firm's innovation threshold (and therefore of $\boldsymbol{Y}$) is shifted leftward by an increase in imitation cost.

$\widetilde{K} < K < \widehat{K}$ **subcase** This is also a war of attrition but since $L(Y)$ is decreasing over $\left(Y^L, Y^F\right)$, increasing over $\left(Y^F, Y^M\right)$ and decreasing over $\left(Y^M, \infty\right)$ its form is nonstandard. To identify the support of the mixed strategies, observe that any $Y_i \in \left(y, Y^L\right)$ is a strictly dominated strategy and no player puts positive probability on Y^L in a symmetric equilibrium. Similarly any $Y_i \in \left(\sup\left\{Z < Y^M \middle| L(Z) > L(Y^M)\right\}, Y^M\right)$ is strictly dominated by Y^M, and no player puts positive probability on Y^M in a symmetric equilibrium. Investment thresholds are therefore continuously distributed over $\left[Y^L, \sup\left\{Z < Y^M \middle| L(Z) > L(Y^M)\right\}\right] \cup [Y^M, \infty)$.

Letting G_b denote the equilibrium distribution, assume that firm $j \neq i$ randomizes her investment trigger. Firm i randomizes if $\partial\left(\int_{Y^M}^{\infty} V(Y_i, s) g_b(s) ds\right)/\partial Y_i = 0$ over the support, that is if G_b satisfies $[F(Y) - L(Y)]\, g_b(Y_i) = -L'(Y)\,[1 - G_b(Y)]$ for all $Y \in \left(Y^L, \sup\left\{Z < Y^M \middle| L(Z) > L(Y^M)\right\}\right)$ and $[F(Y) - M(Y)]\, g_b(Y) = -M'(Y)\,[1 - G_b(Y)]$ for all $Y \in \left(Y^M, \infty\right)$. The former condition holds for

$$\begin{aligned} G_{a'}(Y) &= 1 - \exp\int_{Y^L}^{Y} \frac{L'(s)}{F(Y^F) - L(s)} ds \\ &= \frac{L\left(Y^L\right) - L(Y)}{F(Y^F) - L(Y)} \end{aligned}$$

while the latter condition is satisfied by G_α so that the equilibrium distribution is

$$G_b(Y) = \begin{cases} 0, & Y < Y^L \\ G_{a'}(Y), & Y^L \leq Y \leq \overline{Y} \\ G_{a'}\left(\overline{Y}\right), & \overline{Y} < Y < Y^M \\ G_{a'}\left(\overline{Y}\right) + \left(1 - G_{a'}\left(\overline{Y}\right)\right) G_a(Y), & Y \geq Y^M \end{cases}$$

where $\overline{Y} := \sup\left\{ Z < Y^M \middle| L(Z) > L(Y^M) \right\}$ for compactness. To establish that an increase in imitation cost accelerates innovation as claimed in the text, observe that $L\left(Y^L\right) - L\left(Y\right)$ is independent of K while $F\left(Y^F\right) - L\left(Y\right)$ is decreasing in K so $\partial G_{a'}/\partial K > 0$ and it also holds that $\partial\left(\sup\left\{ Z < Y^M \middle| L(Z) > L(Y^M) \right\}\right)/\partial K > 0$ so that, given that $\partial G_a/\partial K > 0$ as seen above, $\partial G_b/\partial K > 0$ over the relevant range.

A.4.3 Part (ii)

There are two subcases for this part, $\widehat{K} < K < I$ and $K \geq I$ (see Figures 3 and 4). It is simpler to begin with the subcase $K \geq I$ which is standard.

$K \geq I$ subcase For $K \geq I$, $L\left(Y^F\right) \geq F\left(Y^F\right)$ so there exists a unique $Y^P < Y^F$ such that $L\left(Y^P\right) = F\left(Y^F\right)$. The preemption range is $\mathcal{P} = \left(Y^P, Y^F\right)$, and firms seek to innovate just before their rival for any $Y_{-i} \in \mathcal{P}$. In equilibrium both firms set $Y_i = Y^P$ which by the tie-breaking rule results in either firm innovating at Y^P with equal probability.

As joint investment equilibria also arise in preemption models, it must be verified that this is not the case here. Investment at the optimal simultaneous investment threshold Y^M by both firms results in a payoff $M\left(Y^M\right)$ whereas for $K \geq I$ (for $K > \widetilde{K}$ in fact), $L\left(Y^L\right) > M\left(Y^M\right)$, so joint investment cannot be an equilibrium.

$\widehat{K} < K < I$ subcase For $\widehat{K} < K < I$, the condition $L\left(Y\right) = F\left(Y\right)$ has two roots Y^P and $\sup \mathcal{P}$ with $Y^P < Y^L$ and $\sup \mathcal{P} \in \left(Y^L, Y^F\right)$. For a given $Y > \sup \mathcal{P}$, $L\left(Y\right) \leq F\left(Y^F\right)$ so playing beyond the preemption range $\mathcal{P}$ is a dominated strategy. Over $\mathcal{P}$ the firms preempt one another as in the previous subcase and in equilibrium both firms invest at Y^P, which by the tie-breaking rule (7) results in either firm investing at Y^P with equal probability.

A.4.4 Part (iii)

If $K = \widehat{K}$, then $L\left(Y^L\right) = F\left(Y^F\right)$ and $\left(Y^L, Y^L\right)$ is the only symmetric equilibrium.

A.5 Proposition 3 (comparative statics of $\widehat{K}$)

For the comparative static in $\left(\pi^M/\pi^D\right)$ evaluating the relevant partial derivatives and rearranging yields

$$\frac{\partial \widehat{K}}{\partial\left(\pi^M/\pi^D\right)}=-\beta\frac{\frac{\pi^M}{\pi^D}-1}{\frac{\pi^M}{\pi^D}\left(1+\beta\left(\frac{\pi^M}{\pi^D}-1\right)\right)}\widehat{K}$$

so $\partial\widehat{K}/\partial\left(\pi^M/\pi^D\right)<0$ directly, whereas for the comparative statics in β

$$\begin{aligned}\frac{\partial \widehat{K}}{\partial\beta} &= \left(-\ln\widehat{K}+\frac{\left(\left(\frac{\pi^M}{\pi^D}-1\right)\left(\frac{\pi^M}{\pi^D}\right)^{\beta}-\left(1+\beta\left(\frac{\pi^M}{\pi^D}-1\right)\right)\left(\frac{\pi^M}{\pi^D}\right)^{\beta}\ln\frac{\pi^M}{\pi^D}\right)\left(\frac{\pi^M}{\pi^D}\right)^{\beta}}{\left(1+\beta\left(\frac{\pi^M}{\pi^D}-1\right)\right)\left(\frac{\pi^M}{\pi^D}\right)^{2\beta}}\right)\frac{\widehat{K}}{\beta-1} \\ &= -\frac{1}{\beta-1}\left(\ln\widehat{K}-\frac{\frac{\pi^M}{\pi^D}-1}{1+\beta\left(\frac{\pi^M}{\pi^D}-1\right)}\right)\widehat{K}.\end{aligned}$$

The sign of $\partial\widehat{K}/\partial\beta$ is the opposite of that of the (bracketed) middle term. Applying the logarithm inequality $\ln x>(x-1)/x$ for $x>0, x\neq 1$ with $x=\left(1+\beta\left(\left(\pi^M/\pi^D\right)-1\right)\right)/\left(\pi^M/\pi^D\right)$ yields

$$\ln\left(\frac{1+\beta\left(\frac{\pi^M}{\pi^D}-1\right)}{\frac{\pi^M}{\pi^D}}\right)>\frac{(\beta-1)\left(\frac{\pi^M}{\pi^D}-1\right)}{1+\beta\left(\frac{\pi^M}{\pi^D}-1\right)}$$

so $\partial\widehat{K}/\partial\beta<0$ and hence $\partial\widehat{K}/\partial\alpha, \partial\widehat{K}/\partial\sigma>0$.

A.6 Proposition 5 ($K^W\geq\widetilde{K}$)

If $K<\widetilde{K}$, firms randomize investment triggers over $\left[Y^M,\infty\right)$ according to the distribution $G_a(Y)$ and imitator entry is immediate. As discussed in the text, producer surplus is constant over $\left[0,\widetilde{K}\right]$, whereas since $\boldsymbol{Y}$ is stochastically decreasing in K (see Section A.4) consumer surplus from both innovation and imitation is increasing in K over this range.

A.7 Lemma 1 (characterization of $\widehat{Y}_P^W$)

Suppose $K > \widehat{K}$ so in equilibrium the innovation threshold is $\boldsymbol{Y} = Y^P \in \left[(r-\alpha)\, I/\pi^M, Y^L\right)$ and imitation occurs at Y^F. The social welfare function (9) is

$$\begin{aligned} W(K) &= \left(\frac{\pi^M + s^M}{r-\alpha} Y^P - I\right)\left(\frac{y}{Y^P}\right)^{\beta} + \left(\frac{\left(2\pi^D + s^D\right) - \left(\pi^M + s^M\right)}{r-\alpha} Y^F - K\right)\left(\frac{y}{Y^F}\right)^{\beta} \\ &= \left(\frac{\pi^M + s^M}{r-\alpha} Y^P - I\right)\left(\frac{y}{Y^P}\right)^{\beta} + \left(\beta\left(\frac{s^D - \left(\pi^M + s^M\right)}{\pi^D}\right) + \beta + 1\right)\frac{\pi^D y}{\beta\,(r-\alpha)}\left(\frac{y}{Y^F}\right)^{\beta-1} \end{aligned}$$

and the value $W(\widehat{K})$ results by continuity since $\lim_{K \to \widehat{K}} Y^P = Y^L$.

A preliminary step is to obtain an expression for dY^P/dK which is used subsequently in the computation of $W'(K)$. Recall that Y^P is defined implicitly by the condition $L\left(Y^P\right) = F\left(Y^F\right)$. Dividing this identity by y^β and grouping the Y^P and Y^F terms yields a more compact form,

$$\begin{aligned} \frac{\pi^M}{r-\alpha}\frac{1}{\left[Y^P\right]^{\beta-1}} - \frac{I}{\left[Y^P\right]^{\beta}} &= \frac{\pi^M}{r-\alpha}\frac{1}{\left[Y^F\right]^{\beta-1}} - \frac{K}{\left[Y^F\right]^{\beta}} \\ &= \left(\frac{\pi^M}{r-\alpha} - \frac{\beta-1}{\beta}\frac{\pi^D}{r-\alpha}\right)\frac{1}{\left[Y^F\right]^{\beta-1}}. \end{aligned}$$

Observe that $\left[Y^P\right]^{-\beta} = \left[\left(\frac{\pi^M}{r-\alpha} Y^F - K\right) / \left(\frac{\pi^M}{r-\alpha} Y^P - I\right)\right]\left[Y^F\right]^{-\beta}$. The above condition has the form $f_{\mathrm{A.6}}\left(Y^P\right) = g_{\mathrm{A.6}}\left(K\right)$, so $dY^P/dK = g'_{\mathrm{A.6}}(K)/f'_{\mathrm{A.6}}(Y^P)$ with

$$\begin{aligned} f'_{\mathrm{A.6}}\left(Y^P\right) &= -\left(\beta-1\right)\frac{\pi^M}{r-\alpha}\frac{1}{\left[Y^P\right]^{\beta}} + \beta\frac{I}{\left[Y^P\right]^{\beta+1}} \\ &= \left(\beta-1\right)\frac{\pi^M}{r-\alpha}\frac{Y^L - Y^P}{\left[Y^P\right]^{\beta+1}} > 0 \end{aligned}$$

and

$$g'_{\mathrm{A.6}}\left(K\right) = -\frac{\beta-1}{Y^F}\frac{dY^F}{dK} g_{\mathrm{A.6}}\left(K\right) < 0.$$

Substituting $g(K) = f\left(Y^P\right)$ into $g'(K)$ and then developing yields

$$\frac{dY^P}{dK} = -\frac{\beta-1}{K}\frac{f_{\mathrm{A.6}}\left(Y^P\right)}{f'_{\mathrm{A.6}}(Y^P)} = -\frac{r-\alpha}{\pi^M}\frac{\frac{\pi^M}{r-\alpha}Y^P - I}{Y^L - Y^P}\frac{Y^P}{K}.$$

Evaluating $W'(K)$ gives

$$W'\left(K\right) = \left(-\left(\beta-1\right)\frac{s^M + \pi^M}{r-\alpha} Y^P + \beta I\right)\frac{y^\beta}{\left[Y^P\right]^{\beta+1}}\frac{dY^P}{dK} - \left(\beta\left(\frac{s^D - \left(\pi^M + s^M\right)}{\pi^D}\right) + \beta + 1\right)\left(\frac{y}{Y^F}\right)^{\beta}.$$

Substitute the expression for dY^P/dK above to get

$$W'(K) = \left(-(\beta-1)\frac{s^M+\pi^M}{r-\alpha}Y^P+\beta I\right)\left(-\frac{r-\alpha}{\pi^M}\frac{\frac{\pi^M}{r-\alpha}Y^P-I}{Y^L-Y^P}\frac{1}{K}\right)\left(\frac{y}{Y^P}\right)^\beta \\ -\left(\beta\left(\frac{s^D-(\pi^M+s^M)}{\pi^D}\right)+\beta+1\right)\left(\frac{y}{Y^F}\right)^\beta.$$

Substituting for $\left[Y^P\right]^{-\beta}$ in the first term, rearranging, and factoring $\left(y/Y^F\right)^\beta$ yields

$$W'(K) = \left((\beta-1)\left(\left(\frac{s^M}{\pi^M}+1\right)Y^P-Y^L\right)\frac{\frac{\pi^M}{r-\alpha}Y^F-K}{Y^L-Y^P}\frac{1}{K}-\left(\beta\left(\frac{s^D-(\pi^M+s^M)}{\pi^D}\right)+\beta+1\right)\right)\left(\frac{y}{Y^F}\right)^\beta$$

$$= \left(\left((\beta-1)\left(\left(\frac{s^M}{\pi^M}+1\right)\left(Y^P-Y^L\right)+\frac{s^M}{\pi^M}Y^L\right)\right)\frac{\frac{\beta}{\beta-1}\frac{\pi^M}{\pi^D}-1}{Y^L-Y^P} \\ -\left(\beta\left(\frac{s^D-(\pi^M+s^M)}{\pi^D}\right)+\beta+1\right)\right)\left(\frac{y}{Y^F}\right)^\beta$$

$$= \left(-\left(\frac{s^M}{\pi^M}+1\right)\left(\beta\frac{\pi^M}{\pi^D}-(\beta-1)\right)+\frac{s^M}{\pi^M}\left(\beta\frac{\pi^M}{\pi^D}-(\beta-1)\right)\frac{Y^L}{Y^L-Y^P} \\ -\left(\beta\left(\frac{s^D-(\pi^M+s^M)}{\pi^D}\right)+\beta+1\right)\right)\left(\frac{y}{Y^F}\right)^\beta.$$

Regrouping the constant (non-Y) terms gives

$$-\left(\frac{s^M}{\pi^M}+1\right)\left(\beta\frac{\pi^M}{\pi^D}-(\beta-1)\right)-\left(\beta\left(\frac{s^D-(\pi^M+s^M)}{\pi^D}\right)+\beta+1\right)=(\beta-1)\frac{s^M}{\pi^M}-\beta\frac{s^D}{\pi^D}-2.$$

Therefore

$$W'(K) = \left(\left(\beta\frac{s^M}{\pi^D}-(\beta-1)\frac{s^M}{\pi^M}\right)\frac{Y^L}{Y^L-Y^P}+(\beta-1)\frac{s^M}{\pi^M}-\beta\frac{s^D}{\pi^D}-2\right)\left(\frac{y}{Y^F}\right)^\beta.$$

Since $\beta\left(s^M/\pi^D\right)>(\beta-1)\left(s^M/\pi^M\right)$ and $\lim_{K\to\widehat{K}}Y^P=Y^L$, $\lim_{K\to\widehat{K}}W'(K)=+\infty$. Moreover $Y^L/\left(Y^L-Y^P\right)$ and y/Y^F are both decreasing functions of K, so W' is decreasing over $\left(\widehat{K},\infty\right)$. Provided that $\lim_{K\to\infty}W'(K)<0$ therefore, the first-order condition has a unique root in $\left(\widehat{K},\infty\right)$.

Since $\lim_{K\to\infty} Y^P = (r-\alpha) I/\pi^M = ((\beta-1)/\beta) Y^L$, the sign of $\lim_{K\to\infty} W'(K)$ is the same as that of

$$\begin{aligned} & \beta\left(\beta\frac{s^M}{\pi^D} - (\beta-1)\frac{s^M}{\pi^M}\right) + (\beta-1)\frac{s^M}{\pi^M} - \beta\frac{s^D}{\pi^D} - 2 \\ = \quad & \beta^2\frac{s^M}{\pi^D} - (\beta-1)^2\frac{s^M}{\pi^M} - \beta\frac{s^D}{\pi^D} - 2. \end{aligned}$$

Taken as a function of β, this is a quadratic $\Delta(\beta)$, with $\Delta(1) = \left(s^M - s^D - 2\pi^D\right)/\pi^D < 0$ and $\lim_{\beta\to\infty}\Delta(\beta) = \infty$. Therefore there exists a unique $\widehat{\beta} > 1$ such that $\Delta\left(\widehat{\beta}\right) = 0$. It follows that the constrained optimization problem $\max_{K\geq\widehat{K}} W(K)$ has a unique optimum, which is finite (interior) if $\beta < \widehat{\beta}$ and infinite otherwise. Provided that $\beta < \widehat{\beta}$, the first-order condition then yields the expression for $\widehat{Y}_P^W$ given in the proposition.

A.8 Lemma 2 (condition for $\widehat{K}^W \geq \widehat{K}$)

The lemma is established by first deriving an upper bound of W over $\left[0, \widehat{K}\right)$ (attrition) and then comparing this bound with the maximum value over $\left[\widehat{K}, \infty\right]$.

Under attrition, expected social welfare (9) can be expressed

$$\begin{aligned} W(K) = \mathbb{E}_y & \left[\left(\frac{s^M+\pi^M}{r-\alpha}\boldsymbol{Y} - I\right)\left(\frac{y}{\boldsymbol{Y}}\right)^\beta\right] \\ & + \mathbb{E}_y\left[\left(\frac{\left(s^D+2\pi^D\right) - \left(s^M+\pi^M\right)}{r-\alpha}\max\left\{\boldsymbol{Y}, Y^F\right\} - K\right)\left(\frac{y}{\max\left\{\boldsymbol{Y}, Y^F\right\}}\right)^\beta\right]. \end{aligned}$$

The first integrand is a quasiconcave function of investment threshold with a maximum at $(\beta(r-\alpha) I)/((\beta-1)(s^M+\pi^M)) \leq Y^L \leq \boldsymbol{Y}$, which therefore is decreasing over $\left(Y^L, \infty\right)$. Therefore

$$\begin{aligned} \mathbb{E}_y\left[\left(\frac{s^M+\pi^M}{r-\alpha}\boldsymbol{Y} - I\right)\left(\frac{y}{\boldsymbol{Y}}\right)^\beta\right] & \leq \left(\frac{s^M+\pi^M}{r-\alpha}Y^L - I\right)\left(\frac{y}{Y^L}\right)^\beta \\ & \leq \left(\beta\frac{s^M}{\pi^M} + 1\right)\frac{I}{\beta-1}\left(\frac{y}{Y^L}\right)^\beta. \end{aligned} \tag{13}$$

The bound on the second summand uses the assumption that the static entry incentive is excessive,

$$\begin{aligned} \mathbb{E}_y & \left[\left(\frac{\left(s^D+2\pi^D\right) - \left(s^M+\pi^M\right)}{r-\alpha}\max\left\{\boldsymbol{Y}, Y^F\right\} - K\right)\left(\frac{y}{\max\left\{\boldsymbol{Y}, Y^F\right\}}\right)^\beta\right] \\ & \leq \mathbb{E}_y\left[\left(\frac{\pi^D}{r-\alpha}\max\left\{\boldsymbol{Y}, Y^F\right\} - K\right)\left(\frac{y}{\max\left\{\boldsymbol{Y}, Y^F\right\}}\right)^\beta\right]. \end{aligned}$$

The term on the right-hand side is the equilibrium expected payoff of a follower, $\mathbb{E}_y\left[F\left(\max\left\{\boldsymbol{Y}, Y^F\right\}\right)\right]$, that firm i obtains by setting $Y_i = \infty$. Equilibrium payoffs are constant over the support of mixed strategies, so $\mathbb{E}_y\left[F\left(\max\left\{\boldsymbol{Y}, Y^F\right\}\right)\right] = \max\left\{L\left(Y^L\right), L\left(Y^M\right)\right\}$. This last term is maximized for $K = \widehat{K}$, by Proposition 4. Therefore

$$\mathbb{E}_y\left[\left(\frac{\left(s^D + 2\pi^D\right) - \left(s^M + \pi^M\right)}{r-\alpha}\max\left\{\boldsymbol{Y}, Y^F\right\} - K\right)\left(\frac{y}{\max\left\{\boldsymbol{Y}, Y^F\right\}}\right)^{\beta}\right] \leq \frac{\widehat{K}}{\beta - 1}\left(\frac{y}{\widehat{Y}^F}\right)^{\beta}$$

where $\widehat{Y}^F := \left(\beta\left(r-\alpha\right)\widehat{K}\right)/((\beta-1)\pi^D)) = \left(\widehat{K}/I\right)\left(\pi^M/\pi^D\right)Y^L$. Substituting for $\left(\widehat{K}/I\right)^{-(\beta-1)}$ gives

$$\begin{aligned}\mathbb{E}_y\left[\left(\frac{\left(s^D + 2\pi^D\right) - \left(s^M + \pi^M\right)}{r-\alpha}\max\left\{\boldsymbol{Y}, Y^F\right\} - K\right)\left(\frac{y}{\max\left\{\boldsymbol{Y}, Y^F\right\}}\right)^{\beta}\right] &\leq \frac{1}{1+\beta\left(\frac{\pi^M}{\pi^D} - 1\right)}\frac{I}{\beta - 1}\left(\frac{y}{Y^L}\right)^{\beta} \\ &< \frac{0.5I}{\beta-1}\left(\frac{y}{Y^L}\right)^{\beta}.\end{aligned} \tag{14}$$

Combining (13) and (14) yields

$$\max_{K\in\left[0,\widehat{K}\right)} W\left(K\right) < \left(\beta\frac{s^M}{\pi^M} + 1.5\right)\frac{I}{\beta-1}\left(\frac{y}{Y^L}\right)^{\beta}.$$

Over $\left[\widehat{K}, \infty\right]$, social welfare can be evaluated exactly, yielding[28]

$$\max_{K\in\left[\widehat{K},\infty\right]} W\left(K\right) = \begin{cases} \frac{s^M}{\pi^M}\frac{(1+\psi)^{\beta}}{\psi}\frac{I}{\beta-1}\left(\frac{y}{Y^L}\right)^{\beta}, & \beta < \widehat{\beta} \\ \left(\frac{\beta}{\beta-1}\right)^{\beta}\frac{s^M}{\pi^M}I\left(\frac{y}{Y^L}\right)^{\beta}, & \beta \geq \widehat{\beta}\end{cases} \tag{15}$$

where ψ is given in Lemma 1. By revealed preference, the optimal welfare level is at least as large as if the regulator sets an infinite cost of imitation, so

$$\max_{K\in\left[\widehat{K},\infty\right]} W\left(K\right) \geq \left(\frac{\beta}{\beta-1}\right)^{\beta}\frac{s^M}{\pi^M}I\left(\frac{y}{Y^L}\right)^{\beta}.$$

A sufficient condition for $\max_{K\in\left[0,\widehat{K}\right)} W\left(K\right) < \max_{K\in\left[\widehat{K},\infty\right]} W\left(K\right)$ is therefore

$$\left(\frac{\beta}{\beta-1}\right)^{\beta}\frac{s^M}{\pi^M}I\left(\frac{y}{Y^L}\right)^{\beta} > \left(\beta\frac{s^M}{\pi^M} + 1.5\right)\frac{I}{\beta-1}\left(\frac{y}{Y^L}\right)^{\beta}.$$

Cancelling common terms and rearranging yields

$$\left(\left(\frac{\beta}{\beta-1}\right)^{\beta-1} - 1\right)\frac{s^M}{\pi^M} > \frac{3}{2\beta}$$

which establishes the condition in the text.

[28]The derivation is available from the authors.

A.9 Proposition 7 (specific welfare cases)

For part (i), we consider the limiting case $s^M = 0$ (assuming that $s^D > 0$) and then use the continuity of $W(K)$. From the expression of $W(K)$ (9), only producer surplus and the consumer surplus from imitation matter in this case. Both of these are decreasing in K over $\left(\widehat{K}, \infty\right]$ so any maximum of K must lie in $\left(\widetilde{K}, \widehat{K}\right]$ (the lower bound is the one given by Proposition 5). Moreover, since producer surplus is maximized at $\widehat{K}$ (Proposition 4), the sign of the left derivative of welfare at $\widehat{K}$, $\lim_{K\to\widehat{K}^-} W'(K)$, depends only on the consumer surplus from imitation. Given the innovation threshold under attrition $\boldsymbol{Y}$, this term is

$$\begin{aligned}
\frac{\left(s^D - s^M\right) y^\beta}{r-\alpha} \mathbb{E}_y \frac{1}{\left[\max\left\{\boldsymbol{Y}, Y^F\right\}\right]^{(\beta-1)}} &= \frac{\left(s^D - s^M\right) y^\beta}{r-\alpha} \left(\frac{1}{\left[Y^F\right]^{(\beta-1)}} \Pr\left\{\boldsymbol{Y} \le Y^F\right\} + \mathbb{E}_{\boldsymbol{Y} > Y^F} \frac{1}{\boldsymbol{Y}^{(\beta-1)}} \right) \\
&= \frac{\left(s^D - s^M\right) y^\beta}{(r-\alpha)\left[Y^F\right]^{(\beta-1)}} \left(G_c\left(\overline{Y}_L\right) + \int_{Y^M}^{\infty} \left(\frac{Y^F}{s}\right)^{\beta-1} dG_c(s) \right)
\end{aligned}$$

where $G_c = 1 - \left(1 - G_b\left(Y\right)\right)^2$ denotes the equilibrium distribution of $\boldsymbol{Y}$. Since $G_b\left(\overline{Y}\right)\big|_{K=\widehat{K}} = 1$ (recall that $\overline{Y} := \sup\left\{ Z < Y^M \,\middle|\, L(Z) > L(Y^M)\right\}$) $G_c\left(\overline{Y}\right)\big|_{K=\widehat{K}} = 1$ and the right hand term vanishes at $\widehat{K}$ (note that $dG_c/dY = 2\left(1 - G_b\right)\left(dG_b/dY\right)$). Therefore at $\widehat{K}$ only the direct effect $\partial Y^F / \partial K$ remains, hence

$$\lim_{K\to\widehat{K}^-} W'(K) = -(\beta-1)\frac{s^D - s^M}{r-\alpha}\left(\frac{y}{Y^F}\right)^\beta \frac{\partial Y^F}{\partial K} \le 0$$

with strict inequality if $s^D > s^M$. For $s^M = 0$ therefore, $\lim_{K\to\widehat{K}^-} W'(K) < 0$ so the maximum of W lies in $\left(\widetilde{K}, \widehat{K}\right)$.

The argument for part (ii) is straightforward. From the expression of $W(K)$ (9), only producer surplus and the consumer surplus from innovation matter in this case. The first of these is weakly increasing in K over $\left[0, \widehat{K}\right]$ and the second is increasing over $\mathbb{R}_+$, so the maximum of $W(K)$ lies in $\left(\widehat{K}, \infty\right]$.

A.10 Proposition 8 (predetermined sequence vs. endogenous innovation welfare)

For sufficiently large β ($\beta \ge \widehat{\beta}$ and $s^M/\pi^M \ge \Omega(\beta)$), with endogenous firm roles the socially optimal imitation cost $\widehat{K}^W = \infty$ results in a level of welfare

$$\widehat{W}\left(\widehat{K}^W\right) = \left(\frac{\beta}{\beta-1}\right)^\beta \frac{s^M}{\pi^M} I \left(\frac{y}{Y^L}\right)^\beta$$

whereas with predetermined roles, for $\beta \geq \widetilde{\beta}$, $\widetilde{K}^W = \infty$ results in a level of welfare

$$\widetilde{W}^W\left(\widetilde{K}^W\right) = \left(\beta \frac{s^M}{\pi^M} + 1\right) \frac{I}{\beta - 1} \left(\frac{y}{Y^L}\right)^\beta .$$

Then

$$\begin{aligned} \widehat{W}\left(\widehat{K}^W\right) - \widetilde{W}^W\left(\widetilde{K}^W\right) &= \left(\beta\left(\left(\frac{\beta}{\beta-1}\right)^{\beta-1} - 1\right)\frac{s^M}{\pi^M} - 1\right)\frac{I}{\beta-1}\left(\frac{y}{Y^L}\right)^\beta \\ &\geq \frac{0.5I}{\beta-1}\left(\frac{y}{Y^L}\right)^\beta \end{aligned}$$

where the second line uses $s^M/\pi^M \geq \Omega(\beta)$.

A.11 Proposition 10 (takeovers and licensing)

Before establishing the proposition we verify the claim made in the text, *i.e.* that attrition does not occur if $\pi^M/\pi^D \geq \beta + 1$. Consider the limiting case $K = 0$. The leader payoff $L^T(Y)$ is maximized at the threshold $Y^T = (\beta(r-\alpha)I)/((\beta-1)(\pi^M - \pi^D))$. The investment game is (weakly) preemptive if $L^T\left(Y^T\right) \geq F\left(Y^T\right)$, that is if

$$\left(\frac{\pi^M - \pi^D}{r-\alpha} Y^T - I\right)\left(\frac{y}{Y^T}\right)^\beta \geq \frac{\pi^D}{r-\alpha} Y^T \left(\frac{y}{Y^T}\right)^\beta$$

which yields the desired condition on π^M/π^D.

To establish the proposition we first verify that a takeover is the preferred instrument. The condition $L^T(Y) \geq L^L(Y)$ (with $K = K_0$) works out to

$$\left(\frac{\pi^M - \pi^D}{r-\alpha}\max\left\{Y, Y^F\right\} - K_0\right)\left(\frac{y}{\max\{Y, Y^F\}}\right)^\beta \geq \left(\frac{\pi^D}{r-\alpha}\max\left\{Y, Y^F\right\} - K_0 - K_1\right)\left(\frac{y}{\max\{Y, Y^F\}}\right)^\beta$$

which holds because of the efficiency effect $\pi^M > 2\pi^D$.

That takeovers increase firm profit for $K < \widehat{K}$ follows from $L^T(Y) > L(Y)$ and the rent dissipation property of attrition and preemption.

Similarly, licensing (provided $K_1 > 0$) increases firm profit while leaving the timing of imitation unchanged. It therefore remains to verify that licensing results in earlier innovation. Let $\widehat{K}^L < \widehat{K}$ denote the critical threshold that separates attrition and preemption in the presence of licensing, which solves $L^L\left(Y^L\right) = F(Y^F)$ (as licensing only has a level effect on the leader payoff for $Y < Y^F$ the payoff $L^L(Y)$ is maximized at Y^L). For $K \geq \widehat{K}^L$, as $L^L(Y) > L(Y)$ allowing licensing results in innovation at a threshold that is either lower than the preemption threshold without licensing or weakly lower than the previous possible innovation thresholds. Otherwise if $K < \widehat{K}^L$, the industry is in an attrition regime both with and without licensing and the distribution of innovation thresholds shifts left with licensing.

B Dynamic representation of the investment game

To represent the investment game whose normal form is studied in Section 4 in continuous time, assume that the feasible investment strategies of firms are first-hitting times $\tau(Y_i) := \inf\{t \geq 0 \,|Y(t) \geq Y_i\}$. This applies for instance if managerial decisions consist of hurdle rates $\left(\left(Y_i\pi^M/(r-\alpha)\right)/I\right) - 1$ for innovative investment. Then the distributions of investment times are ordered by the investment thresholds Y_i and the investment game is isomorphic that in Fudenberg and Tirole [14]. Their analysis applies verbatim, by defining simple strategies over investment thresholds rather than time.[29]

B.1 Strategies and payoffs

In the dynamic representation of the investment game the continuation payoffs depend on the current state of the stochastic process $y \in \mathbb{R}_+$ and are accordingly denoted $L^y(Y)$, $F^y(Y)$ and $M^y(Y)$.

An *extended* mixed strategy for player $i \in \{1,2\}$ in state y is a pair of real-valued functions $(G_i^y, \alpha_i^y) : [y,\infty) \times [y,\infty) \to [0,1] \times [0,1]$ such that (a) G_i^y is non-decreasing and right-continuous, (b) $\alpha_i^y(Y) > 0 \Rightarrow G_i^y(Y) = 1$, (c) α_i^y is right-differentiable and (d) if $\alpha_i^y(Y) = 0$ and $Y = \inf\{Z \geq Y, \alpha_i^y(Z) > 0\}$ then α_i^y has positive right-derivative at Y.

Let $G_i^{y-}(Y) := \lim_{Z\to Y^-} G_i^y(Z)$ denote the left-hand limit of $G_i^y(Y)$, $a_i^y(Y) = G_i^y(Y) - G_i^{y-}(Y)$ the magnitude of any jump at Y and set $G_i^{y-}(y) = 0$, $i = 1,2$. Let $\mathcal{Y}_i(y) = \infty$ if $\alpha_i^y(Y) = 0$ for all $Y \geq y$ and $\mathcal{Y}_i(y) = \inf\{Z \geq y, \alpha_i^y(Z) > 0\}$ otherwise, so $\mathcal{Y}(y) = \min\{\mathcal{Y}_1(y), \mathcal{Y}_2(y)\}$ denotes the first threshold at which an investment is certain to occur. Finally let

$$\mu_L(u,v) := \frac{u(1-v)}{u+v-uv} \text{ and } \mu_M(u,v) := \frac{uv}{u+v-uv}.$$

[29] Steg and Thijssen [32] study an investment game with closed-loop stopping times strategies and obtain similar equilibrium outcomes. Their framework accounts for the process exiting the attrition region, whereas with first-hitting time strategies firms remain within the attrition region once it has been attained even if the value of the process subsequently exits.

Firm payoffs are

$$V^y\left(\left(G_i^y,\alpha_i^y\right),\left(G_j^y,\alpha_j^y\right)\right)=$$

$$\left[\int_y^{\max\{\mathcal{Y}(y)^-,y\}}\left(L^y\left(s\right)\left(1-G_j^y(s)\right)dG_i^y(s)+F^y\left(s\right)\left(1-G_i^y(s)\right)dG_j^y(s)\right)+\sum_{Z<\mathcal{Y}(y)}a_i^y\left(Z\right)a_j^y\left(Z\right)M^y(Z)\right]$$
$$+\left(1-G_i^{y-}(\mathcal{Y}\left(y\right))\right)\left(1-G_j^{y-}(\mathcal{Y}\left(y\right))\right)W^{\mathcal{Y}(y)}\left(\left(G_i^y,\alpha_i^y\right),\left(G_j^y,\alpha_j^y\right)\right),$$

$i,j\in\{1,2\}$, $i\neq j$ where

$$W^Y\left(\left(G_i^y,\alpha_i^y\right),\left(G_j^y,\alpha_j^y\right)\right)=\frac{a_j^{y-}\left(Y\right)}{1-G_j^{y-}(Y)}\left(\left(1-\alpha_i^y\left(Y\right)\right)F^y\left(Y\right)+\alpha_i^y\left(Y\right)M^y(Y)\right)+\frac{1-G_j^y(Y)}{1-G_j^{y-}(Y)}L^y\left(Y\right)$$

if $\mathcal{Y}_i\left(y\right)<\mathcal{Y}_j\left(y\right)$,

$$=\frac{a_i^{y-}\left(Y\right)}{1-G_i^{y-}(Y)}\left(\left(1-\alpha_j^y\left(Y\right)\right)L^y\left(Y\right)+\alpha_j^y\left(Y\right)M^y(Y)\right)+\frac{1-G_i^y(Y)}{1-G_i^{y-}(Y)}F^y\left(Y\right)$$

if $\mathcal{Y}_i\left(y\right)>\mathcal{Y}_j\left(y\right)$ and

$$=\begin{cases}M^y(Y), & a_i^y\left(Y\right)=a_j^y\left(Y\right)=1\\ \mu_L(a_i^y(Y),a_j^y\left(Y\right))L^y\left(Y\right)+\mu_L^y(a_j^y(Y),a_i^y\left(Y\right))F^y\left(Y\right) & 0<a_i^y\left(Y\right)+a_j^y\left(Y\right)<2\\ +\mu_M^y(a_i^y(Y),a_j^y\left(Y\right))M^y\left(Y\right), & \\ \frac{\left(a_i^y(Y)\right)'L^y(Y)+\left(a_j^y(Y)\right)'F^y(Y)}{\left(a_i^y(Y)\right)'+\left(a_j^y(Y)\right)'}, & a_i^y\left(Y\right)+a_j^y\left(Y\right)=0\end{cases}$$

if $\mathcal{Y}_i\left(y\right)=\mathcal{Y}_j\left(y\right)$.

For given y, a pair of simple strategies $\left(\left(G_1^y,\alpha_1^y\right),\left(G_2^y,\alpha_2^y\right)\right)$ is a Nash equilibrium if $\left(G_i^y,\alpha_i^y\right)$ maximizes $V^y\left(\left(G_i^y,\alpha_i^y\right),\left(G_j^y,\alpha_j^y\right)\right)$, $i,j\in\{1,2\}$, $i\neq j$. A collection of simple strategies $\left(\left(G_i^y\left(Y\right),\alpha_i^y\left(Y\right)\right)\right)_{y>0}$ is consistent if for $y\leq Y\leq Z$, $G_i^y\left(Z\right)=G_i^y\left(Y\right)+\left(1-G_i^y\left(Y\right)\right)G_i^Y\left(Z\right)$ and $\alpha_i^y\left(Z\right)=\alpha_i^Y\left(Z\right)$. The consistent strategies $\left(\left(G_1^y\left(Y\right),\alpha_1^y\left(Y\right)\right)\right)_{y>0}$ and $\left(\left(G_2^y\left(Y\right),\alpha_2^y\left(Y\right)\right)\right)_{y>0}$ are a perfect equilibrium if the simple strategies $\left(G_1^y\left(Y\right),\alpha_1^y\left(Y\right)\right)$ and $\left(G_2^y\left(Y\right),\alpha_2^y\left(Y\right)\right)$ are a Nash equilibrium for every y.

B.2 Equilibrium

In the closure of the attrition range ($K\leq\widehat{K}$) firms do not resort to extended mixed strategies. Equilibrium strategies are therefore obtained from the unconditional strategies $G_a(Y)$ or $G_b(Y)$ (see Section A.4) according to whether $K\leq\widehat{K}$ or $\widetilde{K}<K<\widehat{K}$. Therefore, letting $G_a^y(Y):=\frac{G_a(Y)-G_a(y)}{1-G_a(y)}$ and $G_b^y(Y):=\frac{G_b(Y)-G_b(y)}{1-G_b(y)}$, $\left(G_i^y\left(Y\right),\alpha_i^y\left(Y\right)\right)=\left(G_a^y(Y),0\right)$ and $\left(G_i^y\left(Y\right),\alpha_i^y\left(Y\right)\right)=\left(G_b^y(Y),0\right)$ are symmetric subgame perfect equilibrium strategies in these two subcases.

In the preemption range ($K > \widehat{K}$) the firms do resort to extended mixed strategies and there are two subcases that we consider successively.

B.2.1 $\widehat{K} < K < I$ subcase

This is the case represented in Figure 3 whose key features are that the preemption range (over which $L^y(Y) > F^y(Y)$) is the bounded interval $\left(Y^P, \sup \mathcal{P}\right) \subset \left(Y^P, Y^F\right)$, and that if a threshold beyond this range is reached, firms play a waiting game as $F^y(Y) > L^y(Y)$ for $Y > \sup \mathcal{P}$. In a dynamic representation of the game, subgame perfect equilibrium strategies must account for this possibility.

At any $y > \sup \mathcal{P}$ the payoff to leading lies below the follower payoff. It is not monotonic however, and there exists a unique threshold $\underline{Y} \in \left(\sup \mathcal{P}, Y^F\right)$ such that $L^y\left(\underline{Y}\right) = L^y\left(Y^M\right)$. The leader payoff is decreasing only over $(\sup \mathcal{P}, \underline{Y}) \cup \left(Y^M, \infty\right)$, and it is this range that constitutes the support of mixed strategies. The attrition subgame is then solved similarly to the $\widetilde{K} < K < \widehat{K}$ case in Section A.4 yielding unconditional distributions

$$G_d\left(Y\right) = 1 - \exp \int_{\sup \mathcal{P}}^{Y} \frac{\left[L^y(s)\right]'}{F\left(\max\left\{Y, Y^F\right\}\right) - L^y(s)} ds$$

and

$$G_e\left(Y\right) = \begin{cases} 0, & Y < \sup \mathcal{P} \\ G_d\left(Y\right), & \sup \mathcal{P} \leq Y \leq \underline{Y} \\ G_d\left(\underline{Y}\right), & \underline{Y} < Y < Y^M \\ G_d\left(\underline{Y}\right) + \left(1 - G_d\left(\underline{Y}\right)\right) G_a\left(Y\right), & Y \geq Y^M \end{cases}$$

so that the conditional distribution is $G_e^y(Y) := \frac{G_e(Y) - G_e(y)}{1 - G_e(y)}$.

If y lies in the preemption range, the reasoning is standard and results in firms investing immediately and using the strategy extensions to coordinate simultaneous investment.

Therefore, the symmetric equilibrium strategies are $(G_i^y(Y), \alpha_i^y(Y))$ with

$$G_i^y(Y) = \begin{cases} 0, & Y < Y^P \\ 1, & Y^P \leq Y < \sup \mathcal{P} \\ G_e^y(Y), & Y \geq \sup \mathcal{P} \end{cases},$$

$$\alpha_i^y(Y) = \begin{cases} 0, & Y < Y^P \\ \frac{L^y(Y) - F^y(Y)}{L^y(Y) - M^y(Y)}, & Y^P \leq Y < \sup \mathcal{P} \\ 0, & Y \geq \sup \mathcal{P} \end{cases}$$

for $i \in \{1, 2\}$.

B.2.2 $\widehat{K} \geq I$ subcase

Here $L^y(Y) > F^y(Y)$ over $\left(Y^P, Y^F\right)$ so the investment game is a standard preemption game. A specificity of the investment game studied here is that for $K > I$, M^y lies strictly above F^y over $\left[Y^F, \infty\right)$. Symmetric equilibrium strategies are nevertheless those of a standard real option game, yielding $(G_i^y(Y), \alpha_i^y(Y))$ with

$$
\begin{aligned}
G_i^y(Y) &= \begin{cases} 0, & Y < Y^P \\ 1, & Y \geq Y^P \end{cases}, \\
\alpha_i^y(Y) &= \begin{cases} 0, & Y < Y^P \\ \frac{L^y(Y) - F^y(Y)}{L^y(Y) - M^y(Y)}, & Y^P \leq Y < Y^F \\ 1, & Y \geq Y^F \end{cases}
\end{aligned}
$$

for $i \in \{1, 2\}$.